

```

1 CGCCCGCGGG CTGAGCTCGG CGATCTGGGC CCCAGCGAGG CGGTGGGGCG
51 GGGCGGGGCG GGGCGGGGCG CGCAGCAGGA GCGAGTGGGG CCGCCCGCCG
101 GGCCACGGAC ACTGTCGCCC GGCGCCAGG TTCCCAACAA GGCTACGCAG
151 AAGAACCCCC TTGACTGAAG CAATGGAGGG GGTCCAGCT GTCTGCTGCC
201 AGGATCCTCG GGCAGAGCTG GTAGAACGGG TGGCAGCCAT CGATGTGACT
251 CACTTGAGAG AGGCAGATGG TGGCCCAGAG CTTACTAGAA ACGGTGTGGA
301 CCCCCACCA CGGGCCAGAG CTGCCTCTGT GATCCCTGGC AGTACTTCAA
351 GACTGCTCCC AGCCCGGCTT AGCCTCTCAG CCAGGAAGCT TTCCCTACAG
401 GAGCGGCCAG CAGGAAGCTA TCTGGAGGCG CAGGCTGGGC CTTATGCCAC
451 GGGGCTGCC AGCCACATCT CCCCCGGGC CTGGCGGAGG CCCACCATCG
501 AGTCCACCA CGTGGCCATC TCAGATGCAG AGGACTGCGT GCAGCTGAAC
551 CAGTACAAGC TGCAGAGTGA GATTGGCAAG GGTGCCTACG GTGTGGTGAG
601 GCTGGCTTAC AACGAAAGTG AAGACAGACA CTATGCAATG AAAGTCCTTT
651 CAAAAAGAA GTTACTGAAG CAGTATGGCT TTCCACGTCG CCCTCCCCCG
701 AGAGGTCCC AGGCTGCCCC GGGAGGACCA GCCAAGCAGC TGCTGCCCTT
751 GGAGCGGGTG TACCAGGAGA TTGCCATCCT GAAGAAGCTG GACCACGTGA
801 ATGTGGTCAA ACTGATCGAG GTCCTGGATG ACCCAGCTGA GGACAACCTC
851 TATTTGGTGT TTGACCTCCT GAGAAAGGGG CCGTCATGG AAGTGCCCTG
901 TGACAAGCCC TTCTCGGAGG AGCAAGCTCG CCTCTACCTG CGGGACGTCA
951 TCCTGGGCTT CGAGTACTTG CACTGCCAGA AGATCGTCCA CAGGGACATC
1001 AAGCCATCCA ACCTGCTCCT GGGGGATGAT GGGCACGTGA AGATCGCCGA
1051 CTTTGGCGTC AGCAACCACT TTGAGGGGAA CGACGCTCAG CTGTCCAGCA
1101 CGGCGGGAAC CCCAGCATTG ATGGCCCCCG AGGCCATTTC TGATTCCGGC
1151 CAGAGCTTCA GTGGGAAGGC CTTGGATGTA TGGGCCACTG GCGTCACGTT
1201 GTA CTGCTTT GTCTATGGGA AGTGCCATT CATCGACGAT TTCATCCTGG
1251 CCCTCCACAG GAAGATCAAG AATGAGCCCG TGGTGTTCCT TGAGGAGCCA
1301 GAAATCAGCG AGGAGCTCAA GGACCTGATC CTGAAGATGT TAGACAAGAA
1351 TCCCGAGACG AGAATTGGGG TGCCAGACAT CAAGTTGCAC CTTGGGTGA
1401 CCAAGAACGG GAGGAGCCC CTTCTTTCGG AGGAGGAGCA CTGCAGCGTG
1451 GTGGAGGTGA CAGAGGGGGA GGTTAAGAAC TCAGTCAGGC TCATCCCAG
1501 CTGGACCACG GTGATCCTGG TGAAGTCCAT GCTGAGGAAG CGTTCCTTTG
1551 GGAACCCGTT TGAGCCCCAG GCACGGAGGG AAGAGCGATC CATGTCTGCT
1601 CCAGGAAACC TACTGGTGAA AGAAGGGTTT GGTGAAGGGG GCAAGAGCCC
1651 AGAGCTCCCC GCGTCCAGG AAGACGAGGC TGCATCCTGA GCCCCTGCAT
1701 GCACCCAGGG CCACCCGGCA GCACACTCAT CCGCGCCTC CAGAGGCCCA
1751 CCCCTCATGC AACAGCCGCC CCCGAGGCA GGGGGCTGGG GACTGCAGCC
1801 CCACTCCCGC CCCTCCCCCA TCGTGCTGCA TGACCTCCAC GCACGCACGT
1851 CCAGGGACAG ACTGGAATGT ATGTCATTTG GGGTCTTGGG GGCAGGGCTC
1901 CCACGAGGCC ATCCTCCTCT TCTTGGCCCT CTTTGGCCTG ACCCATTTCTG
1951 TGGGGAAACC GGGTGCCCAT GGAGCCTCAG AAATGCCACC CGGCTGGTTG
2001 GCATGGCCTG GGGCAGGAGG CAGAGGCAGG AGACCAAGAT GGCAGGTGGA
2051 GGCCAGGCTT ACCACAACGG AAGAGACCTC CCGCTGGGGC CGGGCAGGCC
2101 TGGCTCAGCT GCCACAGGCA TATGGTGGAG AGGGGGGTAC CCTGCCACC
2151 TTGGGTGGT GGCACCAGAG CTCTTGCTA TTCAGACGCT (SEQ ID NO:1)

```

FEATURES:

5'UTR: 1 - 172  
 Start Codon: 173  
 Stop Codon: 1688  
 3'UTR: 1691

FIGURE 1A

Homologous proteins:  
Top BLAST Hits

|            |   | Score | E     |
|------------|---|-------|-------|
| gi 1836161 | gb AAB46910.1  (S83194) Ca2+/calmodulin-dependent pr... | 965   | 0.0   |
| gi 2143629 | pir  A57156 Ca2+/calmodulin-dependent protein kinase... | 959   | 0.0   |
| gi 9256525 | ref NP_061371.1  calcium calmodulin dependent protei... | 946   | 0.0   |
| gi 3882295 | dbj BAA34507.1  (AB018330) KIAA0787 protein [Homo sa... | 594   | e-169 |
| gi 4877951 | gb AAD31507.1 AF140507_1 (AF140507) Ca2+/calmodulin-... | 584   | e-166 |
| gi 7446417 | pir  JC5669 Ca2+/calmodulin-dependent protein kinase... | 581   | e-165 |
| gi 5729895 | ref NP_006540.1  calcium/calmodulin-dependent protei... | 577   | e-163 |
| gi 7446362 | pir  T37317 probable Ca2+/calmodulin-dependent prote... | 409   | e-113 |
| gi 7289880 | gb AAF45480.1  (AE002612) CG17698 gene product [Dros... | 343   | 3e-93 |
| gi 3859986 | gb AAC72943.1  (AF091074) unknown [Homo sapiens]        | 249   | 7e-65 |
| gi 1711543 | sp P50526 SSP1_SCHPO SERINE/THREONINE-PROTEIN KINASE... | 231   | 2e-59 |
| gi 5053103 | gb AAD38851.1 AF156028_1 (AF156028) calcium/calmodul... | 203   | 4e-51 |
| gi 6320976 | ref NP_011055.1  DNA polymerase alpha suppressing pr... | 199   | 7e-50 |
| gi 6911862 | emb CAB72162.1  (AL138649) serine/threonine-protein ... | 193   | 3e-48 |
| gi 6321259 | ref NP_011336.1  Ygl179cp >gi 1170647 sp P43637 KGS9... | 188   | 1e-46 |

BLAST to dbEST:

gi|10204347 /dataset=dbest /taxon=96... 555 e-155

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

Expression information from BLAST dbEST hit:

gi|10204347: Eye (retinoblastoma)

Expression information from PCR-based tissue screening panels:

Human Adult Brain

FIGURE 1B

```

1 MEGGPAVCCQ DPRAELVERV AAIDVTHLEE ADGGPEPTRN GVDPPPRARA
51 ASVIPGSTSR LLPARPSLSA RKLSLQERPA GSYLEAQAGP YATGPASHIS
101 PRAWRRPTIE SHHVAISDAE DCVQLNQYKL QSEIGKGAYG VVRLAYNESE
151 DRHYAMKVLK KKKLLKQYGF PRRPPPRGSQ AAQGGPAKQL LPLERVYQEI
201 AILKKLDHMN VVKLIEVLDD PAEDNLYLVF DLLRKGPVME VPCDKPFSEE
251 QARLYLRDVI LGLEYLHCQK IVHRDIKPSN LLLGDDGHVK IADFGVSNQF
301 EGNDALQSST AGTPARMAPE AISDSGQSFS GKALDWATG VTLYCFVYGK
351 CPFIDDFILA LHRKIKNEPV VFPEEPEISE ELKDLILKML DKNPETRIGV
401 PDIKLHPWWT KNGEEPLPSE EEHCSWEVT EGEVKNSVRL IPSWTTVILV
451 KSMRLKRSGF NPFEQARRE ERSMSAPGNL LVKEGFGEGG KSPELPGVQE
501 DEAAS (SEQ ID NO:2)

```

#### FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN\_GLYCOSYLATION  
N-glycosylation site

147-150 NESE

[2] PDOC00004 PS00004 CAMP\_PHOSPHO\_SITE  
CAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 3

|   |         |      |
|---|---------|------|
| 1 | 71-74   | RKLS |
| 2 | 105-108 | RRPT |
| 3 | 455-458 | RKRS |

[3] PDOC00005 PS00005 PKC\_PHOSPHO\_SITE  
Protein kinase C phosphorylation site

Number of matches: 6

|   |         |     |
|---|---------|-----|
| 1 | 58-60   | TSR |
| 2 | 69-71   | SAR |
| 3 | 100-102 | SPR |
| 4 | 160-162 | SKK |
| 5 | 330-332 | SGK |
| 6 | 437-439 | SVR |

[4] PDOC00006 PS00006 CK2\_PHOSPHO\_SITE  
Casein kinase II phosphorylation site

Number of matches: 7

|   |         |      |
|---|---------|------|
| 1 | 26-29   | THLE |
| 2 | 74-77   | SLQE |
| 3 | 82-85   | SYLE |
| 4 | 117-120 | SDAE |
| 5 | 419-422 | SEEE |
| 6 | 425-428 | SVVE |
| 7 | 430-433 | TEGE |

FIGURE 2A

[5] PDOC00008 PS00008 MYRISTYL  
N-myristoylation site

Number of matches: 2

1 178-183 GSQAAQ  
2 326-331 GQSFSG

[6] PDOC00017 PS00017 ATP\_GTP\_A  
ATP/GTP-binding site motif A (P-loop)

485-492 GFGEKGKS

[7] PDOC00100 PS00107 PROTEIN\_KINASE\_ATP  
Protein kinases ATP-binding region signature

134-157 IGKGAYGWRLAYNESEDRHYAMK

[8] PDOC00100 PS00108 PROTEIN\_KINASE\_ST  
Serine/Threonine protein kinases active-site signature

271-283 IVHRDIKPSNLLL

Membrane spanning structure and domains:

| Helix | Begin | End | Score | Certainty |
|-------|-------|-----|-------|-----------|
| 1     | 339   | 359 | 0.946 | Putative  |

BLAST Alignment to Top Hit:

>gi|1836161|gb|AAB46910.1| (S83194) Ca2+/calmodulin-dependent  
protein kinase IV kinase isoform, CaM-kinase  
alpha [rats, brain, Peptide, 505 aa] [Rattus sp.]  
>gi|4512334|dbj|BAA75246.1| (AB023658)  
Ca/calmodulin-dependent protein kinase alpha,  
CaM-kinase alpha [Rattus norvegicus]  
Length = 505

Score = 965 bits (2468), Expect = 0.0

Identities = 472/505 (93%), Positives = 483/505 (95%)

Query: 1 MEGGPAVCCQDPRAELVERVAATDVTHLEADGGPEPTRNGVDPPPRARAASVIPGSTSR 60

ME PAVCCQDPRAELVERVAAT V HLEEA+ GPEP NGVDPPPRARAASVIPGS SR

Sbjct: 1 MERSPAVCCQDPRAELVERVAATSVAHLEEAEEGPEPASNGVDPPPRARAASVIPGSASR 60

Query: 61 LLPARPSLSARKLSLQERPAGSYLEAQAGPYATGPASHISPRAWRRPTIESHHVAISDAE 120

P RPSLSARK SLQERPAGS LEAQ GPY+TGPASH+SPRAWRRPTIESHHVAISD E

Sbjct: 61 PTPVRPSLSARKFSLQERPAGSCLEAQGPYSTGPASHMSPRAWRRPTIESHHVAISDTE 120

Query: 121 DCVQLNQYKLQSEIGKGAYGWRLAYNESEDRHYAMKVLSSKKLLKQYGFPRRPPPRGSQ 180

DCVQLNQYKLQSEIGKGAYGWRLAYNE EDRHYAMKVLSSKKLLKQYGFPRRPPPRGSQ

Sbjct: 121 DCVQLNQYKLQSEIGKGAYGWRLAYNEREDRHYAMKVLSSKKLLKQYGFPRRPPPRGSQ 180

Query: 181 AAQGGPAKQLLPLERVYQETAILKKLDHNVVKLIEVLDDPAEDNLYLVFDLLRKGPVME 240

A QGGPAKQLLPLERVYQETAILKKLDHNVVKLIEVLDDPAEDNLYLVFDLLRKGPVME

Sbjct: 181 APQGGPAKQLLPLERVYQETAILKKLDHNVVKLIEVLDDPAEDNLYLVFDLLRKGPVME 240

FIGURE 2B

Query: 241 VPCDKPFSEEQARLYLRDVLGLLEYLHCQKIVHRDIKPSNLLLGGDGHVKIADFGVSNQF 300  
 VPCDKPF EEQARLYLRD+ILGLLEYLHCQKIVHRDIKPSNLLLGGDGHVKIADFGVSNQF  
 Sbjct: 241 VPCDKPFEEQARLYLRDILGLLEYLHCQKIVHRDIKPSNLLLGGDGHVKIADFGVSNQF 300

Query: 301 EGNDQLSSTAGTPAFMAPEAISDSGQSFSGKALDWATGVTLYCFVYGKCPFIDDFILA 360  
 EGNDQLSSTAGTPAFMAPEAISD+GQSFSGKALDWATGVTLYCFVYGKCPFID++ILA  
 Sbjct: 301 EGNDQLSSTAGTPAFMAPEAISDTGQSFSGKALDWATGVTLYCFVYGKCPFIDEYILA 360

Query: 361 LHRKIKNEPWFPEEPEISEELKDLILKMLDKNPETRIGVPDIKLHPWTKNGEELPSE 420  
 LHRKIKNE WFPPEEPE+SEELKDLILKMLDKNPETRIGV DIKLHPWTK+GEEPLPSE  
 Sbjct: 361 LHRKIKNEAWFPPEEPEVSEELKDLILKMLDKNPETRIGVSDIKLHPWTKHGEEPLPSE 420

Query: 421 EEHCSWEVTEGEVKNSVRLIPSWTTVILVKSMLRKRSFGNPFEPQARREERSMSAPGNL 480  
 EEHCSWEVTE EVKNSV+LIPSWTTVILVKSMLRKRSFGNPFEPQARREERSMSAPGNL  
 Sbjct: 421 EEHCSWEVTEEEVKNSVKLIPSWTTVILVKSMLRKRSFGNPFEPQARREERSMSAPGNL 480

Query: 481 LVKEGFGEGGKSPELPGVQEDEAAS 505  
 L+KEG GEGGKSPELPGVQEDEAAS  
 Sbjct: 481 LLKEGCGEGGKSPELPGVQEDEAAS 505 (SEQ ID NO:4)

# Hmmer search results (Pfam):

| Model   | Description                                | Score  | E-value | N |
|---------|--|--------|---------|---|
| PF00069 | Eukaryotic protein kinase domain           | 275.1  | 8.8e-79 | 1 |
| CE00022 | CE00022 MAGUK_subfamily_d                  | 45.1   | 3.2e-13 | 1 |
| CE00359 | E00359 bone_morphogenetic_protein_receptor | 26.1   | 1.1e-06 | 1 |
| CE00031 | CE00031 VEGFR                              | 13.6   | 0.00033 | 1 |
| CE00203 | CE00203 ERBB_RECEPTOR                      | 6.7    | 0.16    | 1 |
| CE00334 | E00334 urotrophin_receptor                 | 6.6    | 0.047   | 1 |
| CE00292 | CE00292 PTK_membrane_span                  | -66.6  | 4.1e-05 | 1 |
| CE00287 | CE00287 PTK_Eph_orphan_receptor            | -81.5  | 0.0049  | 1 |
| CE00291 | CE00291 PTK_fgf_receptor                   | -95.1  | 0.0027  | 1 |
| CE00286 | E00286 PTK_EGF_receptor                    | -119.5 | 0.00094 | 1 |
| CE00290 | CE00290 PTK_Trk_family                     | -132.1 | 2.8e-06 | 1 |
| CE00016 | CE00016 GSK_glycogen_synthase_kinase       | -206.4 | 3.1e-05 | 1 |
| CE00288 | CE00288 PTK_Insulin_receptor               | -225.0 | 0.18    | 1 |

# Parsed for domains:

| Model   | Domain | seq-f | seq-t  | hmm-f | hmm-t   | score  | E-value |
|---------|--------|-------|--------|-------|---------|--------|---------|
| CE00031 | 1/1    | 243   | 296 .. | 1039  | 1092 .. | 13.6   | 0.00033 |
| CE00334 | 1/1    | 264   | 297 .. | 670   | 703 ..  | 6.6    | 0.047   |
| CE00203 | 1/1    | 262   | 307 .. | 852   | 897 ..  | 6.7    | 0.16    |
| CE00359 | 1/1    | 271   | 320 .. | 272   | 325 ..  | 26.1   | 1.1e-06 |
| CE00288 | 1/1    | 179   | 382 .. | 1     | 269 []  | -225.0 | 0.18    |
| CE00291 | 1/1    | 128   | 382 .. | 1     | 285 []  | -95.1  | 0.0027  |
| CE00290 | 1/1    | 129   | 390 .. | 1     | 282 []  | -132.1 | 2.8e-06 |
| CE00286 | 1/1    | 128   | 403 .. | 1     | 263 []  | -119.5 | 0.00094 |
| CE00292 | 1/1    | 128   | 405 .. | 1     | 288 []  | -66.6  | 4.1e-05 |
| CE00287 | 1/1    | 128   | 407 .. | 1     | 260 []  | -81.5  | 0.0049  |
| CE00022 | 1/1    | 247   | 409 .. | 118   | 283 ..  | 45.1   | 3.2e-13 |
| PF00069 | 1/1    | 128   | 409 .. | 1     | 278 []  | 275.1  | 8.8e-79 |
| CE00016 | 1/1    | 64    | 481 .. | 1     | 433 []  | -206.4 | 3.1e-05 |

FIGURE 2C

|      |             |            |             |             |            |
|------|-------------|------------|-------------|-------------|------------|
| 1    | CCGCCCCGCGC | ATCCATCTGG | GCCTCAGCGT  | GTCCCCGAGCA | ATCACAACAG |
| 51   | CAGCCGCACA  | ACAACAATC  | ACTTTTACGG  | CCTCCTTAGT  | GGCAGGCACT |
| 101  | GTTCTGAGCG  | CCTTACGGGC | GTTCCCTCCT  | CAGCATCTCA  | CCACGTGCGG |
| 151  | TGAGGTGAGG  | CCCGCTAGAA | CCCCATCTTG  | CGGGCGAGGA  | AAACCCAAGG |
| 201  | CACAGAGGCG  | AAGCCACCTG | CTCAGGGGCT  | CCCAGCCAGG  | AAAGGGTGCA |
| 251  | GCCTGGCTGC  | CTGGCTTCAG | AGCCTGGGCG  | CCAAACCGGG  | TAACAGGGCT |
| 301  | CAGGCTGGAA  | CAGGAAACCT | TCTGCCCCGA  | CTTGCTGGGT  | GACCCCGGGC |
| 351  | CCATCCCCAC  | CCGCTGGGCC | TCCCTCTACC  | TATCTAAGAA  | AAGCAGGGAA |
| 401  | AGGTGTTCAA  | GGGTAAAGGA | GGATGGCCTC  | TTGCTGGAAT  | GGCAACCTCA |
| 451  | AGGAAATACG  | CAAATTTTAT | GGGCCCCGGC  | AGCCTGTGGC  | TTCTGCCTGT |
| 501  | GGCGGCTCTG  | AGTCCCGTAG | TCCCTGCCTA  | GGGCCAAAAA  | GCAGGAGCTC |
| 551  | CTGACTCTGG  | AGTTCATTCT | GTTATATGTG  | CTGGGGCCTG  | AGGCTTGCTG |
| 601  | GGGTTGCCTC  | TCTGAGGCTG | CTTTCTCATC  | TGTCTAATGG  | GGACAGGGCT |
| 651  | GTAACGATCA  | CTATGGCAAC | CACTCATTTA  | TTCAACAAAT  | ATTTATCGAG |
| 701  | TTCCTATCAC  | ATGCCAGGCA | CTGATGATCT  | TTTGGAGACA  | AGGCAGATGA |
| 751  | GCGTCCTAAT  | CTCATGAAAC | TTACATTTCG  | GAGGGAAAAA  | CAAGGCATGC |
| 801  | GGAGTGAGGG  | GAAGGGGCGG | AGGGGTGGGC  | CACCTGCTGG  | GAGGAGCCTG |
| 851  | GCGGGTCCTG  | GAGGGTGTTT | CCAGCTTTGG  | CTTCCTCCTT  | CCTATGCTGT |
| 901  | CTGGTTTCCA  | AGCTCTCCCC | GAAGCTCCAG  | CCCCACTCAC  | TGTCCCTCTC |
| 951  | ACCTCCTCCA  | GGGAGGCCTC | CCTATGCCAC  | AGCCTCTCAC  | CTCCTCTGGG |
| 1001 | GAGGCCTCCT  | TATGCCACAG | CCCCACTCTC  | TGTCCTCTCT  | CACCTCCTCC |
| 1051 | AGGGAGGCCT  | CCCTGTGCCA | CAGCCCCACT  | CCCTGTCCCC  | TCTCACCTCC |
| 1101 | TCCAGGGAGG  | CCTCCTTGTT | CCACAGCCGC  | ACTCACTGTC  | TCCTGCCCTC |
| 1151 | TCTTCCAGGG  | AGGCCTCCCT | GATACTCTAG  | CCTCACTCAG  | CCTCCTCACC |
| 1201 | TCCTTCACTT  | CCTCCAGGGA | GGCCTCCTTG  | ATGTTCCAGC  | CTCATTAAC  |
| 1251 | CCCTCTCACT  | CCTCTGGGTC | CAGCTTCCAT  | GACTTTTCCT  | GTTCTAGTG  |
| 1301 | TGGAGCCTCT  | TCTCTTCCTT | TCTCCATGTC  | AGCACCAGCC  | CCACCGCCTC |
| 1351 | CAGGCTTCTA  | CTCATTCAAC | ACACTGCGTA  | CCGGGCACAG  | GGGGTCTGGA |
| 1401 | CCTCACCTTT  | ACCCTCAGTC | TACCTCCAAA  | CCCTGCTGTG  | AGCCTGGAAA |
| 1451 | ATATGGGAAG  | GCAGGGAATC | CACAGGACAA  | GTCGGGAGAC  | TGGGGCTCAG |
| 1501 | AGTCGGGAAG  | GAGCTGGTCT | AGGGCCCCCT  | GTGGGTCAGC  | AGGCAGGACT |
| 1551 | GGAACCCAGT  | CCTGGCTCCT | CAGTGGCCGG  | TGGACTCCAG  | CCAGCCCTGC |
| 1601 | CTCGTGACA   | TCTGTCAAAG | CAAGGGGATG  | GGGAACGAGC  | GGTAGAGCAG |
| 1651 | GCGCTTCACT  | ATGCGTACTC | TGGGTCTCCC  | TGAGACCCAT  | GTTCTCAGTT |
| 1701 | GCTGTGTGGG  | TTCGGAGGAA | GTTACCAGCA  | GACAGGAAAG  | ATGGAGGGTC |
| 1751 | AGGAGTTCAC  | TCACTTCCTT | CTCCTGAGAA  | CATGCAGAGT  | CCAGCGCAAG |
| 1801 | CAGGGGGAAG  | GGCATCAGGT | TGGGCATGGC  | CAGCGCTCTA  | CAAGCCTGGG |
| 1851 | ACAGAGATGG  | GGGTCTCAGG | CTGAGTGTCA  | GGGTTCACTC  | CGGGGTGAGG |
| 1901 | ATGTAGCCCA  | GGGTATGGC  | TGAAGGTGAG  | GGCTGGGGGT  | CACCTCCCTG |
| 1951 | ATGTTTCAGC  | GCCACACAG  | TGAGTTTGAG  | AACATGAGTC  | TCAGGGGATG |
| 2001 | TCATGCCCTT  | GTTTCACCCC | TCATTCCCCCT | CATTCCCATC  | CCCTTGCTTT |
| 2051 | TTTTTGAAAC  | CGAGTCTTGC | TCCATCACCC  | AGGCTGGAGT  | GTAGTGGCGT |
| 2101 | GATCTTGGCT  | CACTGCAACC | TCCACCTCCC  | AAGTTCACAC  | GATTCTCTG  |
| 2151 | CCTCAGCCTC  | CCGAGTAGAT | GGGATTTTCA  | GTGCACGCCA  | CCATGCCTGG |
| 2201 | CTAATTTTTG  | TATTTTTAAT | AGAGACAGAG  | TTTTGCCATG  | TTAGCCAGGC |
| 2251 | TAGTCTCGAA  | CTTCTGACCT | CAGGTGATCC  | ACCTGCCTCG  | GCCTCCCAAA |
| 2301 | GTGCTGGGAT  | TACAAGTGTG | AGCCACCATG  | TGGGGCCCAT  | CCCCTTGTTT |
| 2351 | TGACAGACGT  | CAATGAGGCA | GGGTGGCTG   | GAGTCGGGAG  | CCCCAGGGAA |
| 2401 | GTCTTCTTGG  | AAGCAGTGAG | AGGGATGGGG  | GTAGGAGGCT  | GAAACATCAA |
| 2451 | GGAGGGCTCC  | CTGGAGGAGG | CGGGTGGGTC  | TGAAGCATCA  | GCAAGGCTTC |
| 2501 | TGAGTTACTA  | GTGTCTAGCT | CAGCTTCCAG  | GAGGCAGTGT  | CGGAGTGCTC |
| 2551 | TGCTGTCAAG  | GGTTGGGACT | CATGACTCAC  | AGGGCTGCAT  | GCTGTGCTGG |
| 2601 | GGCTGAGCTG  | ACCCTGGGCT | CTGCCCCTTC  | CAGTGCTGCT  | GGGCCTCCAG |
| 2651 | GCTTCTGCCC  | TGTCTGTCTT | GATTCCAGAA  | TATCAGATTC  | TCTCTGCTTC |
| 2701 | CCTGTGAAGC  | CAGCAGGCAG | AAGTGACTGC  | CTCTGTTACC  | GGCAGGGATA |
| 2751 | CTGAGGCCCTA | GAGGGCTGGC | ATGCGGCAGA  | ACCGATGTGA  | ATTCATTGAG |
| 2801 | GTCAATAGGA  | CAGACTTGAG | TTTGGGTGTT  | GGCAATCCCG  | GTAAGAGGAA |
| 2851 | CAGCCAGGGC  | AAAGGCATGG | AGGTGGGACC  | CACAGCGCTG  | TGGCTACCTT |
| 2901 | ACCTGGTAGC  | CAGCCTGACA | CCCAGGAGTG  | AAGCCTTCTC  | TGCCTTCTTT |

FIGURE 3A

|      |            |             |            |             |            |
|------|------------|-------------|------------|-------------|------------|
| 2951 | TCTCAGGTTT | CCAACAAGGC  | TACGCAGAAG | AACCCCTTG   | ACTGAAGCAA |
| 3001 | TGGAGGGGG  | TCCAGCTGTC  | TGCTGCCAGG | ATCCTCGGG   | AGAGCTGGTA |
| 3051 | GAACGGGTGG | CAGCCATCGA  | TGTGACTCAC | TTGGAGGAGG  | CAGATGGTGG |
| 3101 | CCCAGAGCCT | ACTAGAAACG  | GTGTGGACCC | CCCACCACGG  | GCCAGAGCTG |
| 3151 | CCTCTGTGAT | CCCTGGCAGT  | ACTTCAAGAC | TGCTCCCAGC  | CCGGCCTAGC |
| 3201 | CTCTCAGCCA | GGAAGCTTTC  | CCTACAGGAG | CGGCCAGCAG  | GAAGCTATCT |
| 3251 | GGAGGCGCAG | GCTGGGCCTT  | ATGCCACGGG | GCCTGCCAGC  | CACATCTCCC |
| 3301 | CCCGGGCCTG | GCGGAGGCC   | ACCATCGAGT | CCCACCACGT  | GGCCATCTCA |
| 3351 | GATGCAGAGG | TTGGTGGGG   | AGAACGAGGG | GTTGTTTCATG | AGCCCCTCAG |
| 3401 | TAGTCTGCAA | TGAAGACTCT  | TTCCTGCCCC | TGTCTGTGCC  | ACACGGCTAT |
| 3451 | CTAGCTTTGG | TTTGCATACC  | CTCAGAGCTG | GGGAGATCAC  | TACCTAACAA |
| 3501 | TATAGCTTCT | TCCCAACCAG  | GGGAGCTCCA | GCTGAGCCAA  | AGGCTGCCTT |
| 3551 | CCCTAAGTCC | TGCTATTCCC  | ACTCCCAGCC | CAGGCCTAGG  | AAATAGGTCT |
| 3601 | CTCCCTGGTC | CCCTATGTAG  | TCTTCTTAGA | GATGTGAAGA  | TAGATGCTAT |
| 3651 | GTCCCCCTTC | CCCCCTAACT  | CTTCTCCAGC | TTGCACCCCT  | CGCCTCTAAT |
| 3701 | TCTGCCTCTT | AGAGTCTGCT  | GTGACTCAGA | AGCGGCCGGC  | CTGCCTCCAG |
| 3751 | CCTCTGGGCT | TCTGCTGGAG  | TTCTTGCCAT | TTAGGTCTGA  | AAGTGAAGTC |
| 3801 | AGGTTCCAAG | CAGTCTACAG  | ATGTCAGGGG | CTGAGCTTTC  | TGTGCCTGAA |
| 3851 | CCCAGGCTCT | CAGCCTCTGT  | GCCCAGGGCT | CCTCATCTTG  | TCCTTGGAGT |
| 3901 | CTAGACCTTC | TCATTGAGCT  | GCTTCTGGAA | ATAGTTGCTC  | ATGGGTTTCT |
| 3951 | CATGGATTAG | GGTCTTCCAG  | ACTCCAGAAT | CCAGACAGGA  | ATTAGCGTTT |
| 4001 | TCCCTTACC  | ACTGCTTCTG  | GGGAACAAGG | CACAGCCATG  | GCGTCACCAT |
| 4051 | CCATGTTTTT | AAACATGAGC  | CACGTCTTCT | CGTCACATAC  | GGGGGCGATG |
| 4101 | GCACCACCAA | CTTCCCCATC  | CAAACTCAAA | AGCTTGGTGA  | GACCTGGGGG |
| 4151 | TCCGGGAATG | AGGAGCTTAT  | GGCCAGAATT | GGACCTGAA   | CGGGCTCTGA |
| 4201 | GGTAGGAGCA | GTGCTGCCTC  | CGGACCCAGC | TCCACCTGGT  | GCTCGCTCTT |
| 4251 | CCCCACAGG  | ACTGCGTGCA  | GCTGAACCAG | TACAAGCTGC  | AGAGTGAGAT |
| 4301 | TGGCAAGGTA | GGAGTGGGCA  | GGCCGAGAGC | AGTGGGGGCT  | TCGGGATTCT |
| 4351 | CTGTTTTGGC | CTGCTCCTTC  | TCTCGTGTGG | GAGGGAACGG  | GAGGCAGAGC |
| 4401 | CAGGCAAGTC | CTAGCCTGGA  | GGTGAGGACA | GTTTCGTGCC  | CTGTGGGAAG |
| 4451 | TACCCAGGTA | CCCAGGGGGA  | GGGTGGAAGA | TGGCTCCTGA  | TTCCCGACTC |
| 4501 | TCTGAGTTCT | TGACAGTGGA  | CAAGGAGGGA | CTGAGGGAGG  | CATGGAGCCA |
| 4551 | TGTGGAGCCA | AGCAGGGGCA  | GTTACCAGGG | CGCAGGAGTC  | CCCTCCCCAT |
| 4601 | CTGCTACAAT | ATTTGCCCGT  | GAGCCAGCTG | GTGGTGGGTA  | GTGCAGATGG |
| 4651 | GGTGCAAGAG | AGACCAGAGC  | TGCTCGGCTC | CCCACCTCCT  | GAGCTGGTCC |
| 4701 | TGGGAGGGGT | TGCCCTGTCC  | AGGTGGGGCT | GACTGATGCC  | TATCTGCAGG |
| 4751 | GTGCCTACGG | TGTGGTGAGG  | CTGGCCTACA | ACGAAAGTGA  | AGACAGACAC |
| 4801 | TATGTGAGTC | TGGGGATACG  | AGGGAGGTGT | TGCCCAAGCC  | AGGCCCTGGA |
| 4851 | AGCCTGAGGG | TGGGGACAGG  | AGTTGTGCTT | AGGAGATAGA  | GGACAGGGCT |
| 4901 | GCCTGAGAGT | GAGCTCCCTG  | TCCCTAGGGG | TATGCAAAGG  | AATGAGCTTC |
| 4951 | CTAACCTTGG | GGATATGCAA  | GCAGAGACTG | GATTCTCTG   | AGGGGAAAGC |
| 5001 | TCCAGAAAGG | CTTGCTGGGG  | GAATAAGGGG | AAGGGCTAGG  | CTCAGATATG |
| 5051 | GCCACCCCA  | ACCCCGCTTA  | ACACTTACCT | GGGCCACACC  | CTCAGGGCCA |
| 5101 | GTAGCAGATG | TCCAGTGTGC  | CTCTCCGGAC | CTCAGTCCAC  | ATGTACCAGC |
| 5151 | CTGTTCTAGC | CCCTGGTGGC  | TGCACAGTAG | TGACATTTCT  | GTCCCTCCTT |
| 5201 | CCTTAGGCAA | TGAAAGTCCT  | TTCCAAAAAG | AAGTTACTGA  | AGCAGTATGG |
| 5251 | CTTTCCACGT | ATGTATCTTC  | TGATCCTGTC | CCTGGGAGCT  | CCTAGCCTGG |
| 5301 | AGGCAGAGGA | GGAGACCTCG  | ATCCTGAGCT | AGTTTTGGCT  | AGGAATGGGG |
| 5351 | TAGAGAGGGA | GACAGCGTGA  | GCAGAGGCCT | GGGGACAGAA  | TGTGCCCTGT |
| 5401 | GGGTTGGGAC | AAGACCACGG  | GCATGCAAGA | CTCTTGCTTG  | AGACTGGTTT |
| 5451 | GGGGGCCACG | GTGAGGCCCA  | GCCACCTGGA | ACAGGTGTTT  | GAGTTCTCTT |
| 5501 | CCTGGTCACA | GGTCGCCCTC  | CCCCGAGAGG | GTCCCAGGCT  | GCCCAGGGAG |
| 5551 | GACCAGCCAA | GCAGCTGCTG  | CCCCTGGAGC | GGGTGTACCA  | GGAGATTGCC |
| 5601 | ATCCTGAAGA | AGCTGGACCA  | CGTGAATGTG | GTCAAAGTGA  | TCGAGGTAGG |
| 5651 | GGGTGGTGGT | GAGCAGGTGG  | GAACCAGCAC | CTGAGTCTCA  | TGGGAGCCGC |
| 5701 | TTCTGGTGCT | GGGGAGCCCC  | TAGCACAGAC | CCAGGGATCT  | TGCCCAGGTG |
| 5751 | GCAGATGTGG | CTGAGGCCCTC | TGAGGACAGG | GCCAGACTTG  | GGGTGGGGCT |
| 5801 | GCAGGAAGGC | TTTGGGGGCC  | CAGCCTGGTC | AGGGATGTTT  | CCAAGTTCCC |
| 5851 | ATGGAGGGTG | AGGGGCTGCC  | CCAGAGGCAA | GAAGTGAGCC  | CCTCATTGCA |

FIGURE 3B

5901 GCTGGAGGGG AGGAAGGCTG GATGTCGTGT GGCGGGCCAG GTTGGGGGTC  
5951 GGTGACTTCT GAGGCCCCAT CAGTCTGGCA CCACCTGTAC ACTTCCTGCT  
6001 TCCTTGCTCG GGGTGGTTGC ATGCATACTA AGGGTTCTGG GGCTGGCAAG  
6051 GACCAGGAGG CCTGGGACCT CCAACCCAC GCCTCCTCAA GCCCCACCCC  
6101 CATGTCTGCT CCCTCTGACC AGGTCTGGA TGACCCAGCT GAGGACAACC  
6151 TCTATTTGGG TGAGTGACCT GGCTCATTCC CACAGCAGCT CACTCAGGGC  
6201 TGGCCCAAGG GCTCCCTTGG GACATGTATG ACCTTCAGGT GGGCGGTGTA  
6251 AATGCACTGA CCTCCTGGGG ACAGAAGAAA AACACACGTT CTGAAGCCCT  
6301 GGATTCCCTT GCCCAGCCCT GCAGAACCAG GCCCAGAATA TCCAGTTAGA  
6351 TTCAACAAAT ATCGCCAAGC CCCACTCCCT GCTTCCCTCT GAGCAGCAAG  
6401 ACAGTGGATC CACGTGGGCT GCGCGCTCAG GTAGATGCAG GAAGCAGGCT  
6451 GCATGGGTTT CCAGACACTG TAGCTCTGTG CCTCAGTTTT CCCACCTATA  
6501 AAACAGGGAT ACTAGTGGTG TCTACCTCAT AGGGTTCCTG TGAAGAGTAA  
6551 ATGAGTAATT ATATGTAAG CACATTCGTT ATTATCCTTG TTAATAGTAA  
6601 TGTTATTATT TTAGTTCCTT GTGTCTGGTT CAGGGCTGGG CTTAGAGGAG  
6651 GCCTCAGAAA ATGGGGCAGA AGAAGAACTG GCTTAGGAAT TAGAGGCTGA  
6701 GGCTTTAGTC TCCACTCCCT ACCCTACCTG CCTGTCTGCT ATGACCTTTA  
6751 GGAAAATTTT TGCCCTTCTT CTGTGCCTCA GTTCCCCCT CTGTAAAAGG  
6801 GCCCCATGCT GATGCTGATG GTTCTCACCT GGCACCTGAG GATCAGATGA  
6851 GACAGGTCCA TAGCAGACCC CACTCTCATG CATTTATTTG CTCTCATATC  
6901 CCAGGTGCC CTGTCTGTG CCTGCCTCGA GTATGCCTGC ATGCCTGCCC  
6951 CCTCTCTAC CCTCCAGAAC AGGGAGGGAC CTTGGCATCG GCTGCTTTGC  
7001 CAGCCAGCTA CACCTTACCT TCTTGTCTTT TCTTTCAGTG TTTGACCTCC  
7051 TGAGAAAAGG GTGAGTTCCC CGTCCTGATC AGGCAGGTCA ATTCTCATCC  
7101 AGGCCTTCTT TCCTTCCCT CCCTGTGTCC CCAGCCCAGG GGTGAGCTAC  
7151 TCTAGGAGAA GTCAGAGACG GAGGCCCTGC CTTAGGGGT AAATAAGAGA  
7201 CCAAGAGGAC CATTCTTTGA AGGCTGATGG GGGTCAGTGA GGCTGAAATA  
7251 GTCAGGGAGA CCTCTGAAA AGGGGACGGA TTTTGACCCA GGCCTTGAAG  
7301 AACTAGGAAG ATAGGGATGG AGGAGAGGGG GAAGAAAGGA GTGTTTTTTA  
7351 GGTAAAAGTA TATAGAGGTG GGACTCAACT CTTACCGTA TTCAAATCAC  
7401 AAAGGTTTT TCAGCTTTCC AACAACTCTG TGAATGGAGT GGGTGGGATT  
7451 CCAGTTGCTC CCATTTGTGA GAGGGAAAGC TAAGGACCAG AGAAGGTACG  
7501 TGGCTTGCTC AAGGTCACAC AGCAAGTCAC TGATGGAGCC CAGGCTTCCA  
7551 CATGTCTGCC CTATGCGGCT TTTGAGGGTA TTTACAGAGC AGATGACATG  
7601 GAGTAATGAG CACGGGGCTG GGTGGTCCGG GACCCTCACT GCCAAGGCTT  
7651 GAATGCAGCC TGCGGCTTGT CCTTTGCCT GGGCGGCTCC CTACAGACCA  
7701 ATCTGGGGAG AGGGGCAGGG AGTGGTGTCC CTTTAAGACT TGGAGGCTTT  
7751 CAAATGTTTT GACCTCTATC CAAAACAAGA AATATATATT TCTATTGCTA  
7801 TCCATATCTG TAATGAAAC CAAAATTTTA CAAAGCAGCA TATATCTTTA  
7851 CTACATGCAA TATATTCTGA TATATTCTAC TTATTTAGGA AAAAAAAAAA  
7901 AAAGCAGTTG CCACCCACTA AATTGATTTT ATGATCCTCT CTTGGGTCTG  
7951 GATCCACGGT TTGAAACAGT GCTCTAAATG GCATCTTTGC AATTGATTAT  
8001 GGACAATTAA GTACTTAGAA GAAGGAATAT CAAGCCAATC AGAAATTAAG  
8051 AGAAAGCTGA TTTGAAATTA TGATTGAAAT GGGATATGTA TGAGTATGTG  
8101 TGCTTTAAGT TTTTTATTAT GTAGCAGAAA AAGCTAATAT CTTGAGTTGT  
8151 AGGGAICTAT GTGGGCACAG GTTTCCCGGG ACGTCCCGAC CACCTGAATG  
8201 GCCGGGTGCC CTGATTTTCT GAATGCCC CTCCCCGAT CTTTCTCCAT  
8251 AGGCCCGTCA TGGAAAGTGC CTGTGACAAG CCCTTCTCGG AGGAGCAAGC  
8301 TCGCTCTAC CTCGGGACG TCATCCTGGG CCTCGAGTAC TGTGAGTGCG  
8351 GGGCAGCTTG CCGACTGGGG CTGGGGCTAG GGGATCTGGC AGGCGGCAGA  
8401 GCCCAGGCTG AGCAGACTCT GAGCAGCTCC CGTCAGTCAG AGCTGACCTG  
8451 CCAATCAGCT TCAGTGGGAG TGGGGCATGC ACGTGTGGCG GGGCCAAAGG  
8501 CCTTTTTGTG GGGTGGGGCG GGCGGTGGAC TCCACTGGGC ATGTGCCAGA  
8551 TCCTTCGTGG TGTCTGGTCC TGTGGGTCTG AGTCCTGGCT GTTCTGTATC  
8601 TTTCTTCTGC TGAGTTCTTA GCCTAGCTTA GCGTTGCCAC GGGGCTTCAA  
8651 GAGATGTGGG AAGGAAGGGA TTTATGTCCA GCTGCTGGGG AGAGTCTGTC  
8701 CTGGCATGGG GCCGGGGCAT GGTGGCAGGG TGGATTTACC TGTGAGGGGC  
8751 CCTAGTCTGA TAAGAGCTCA GGAGGGTGAT GTGAGCTTGG CCTCTGTCTC  
8801 ATTCATTCA TTAGCTACAT TCACTTGCTT GGGGCGATAG GGGTGAAAGA

FIGURE 3C



|       |            |             |            |            |             |
|-------|------------|-------------|------------|------------|-------------|
| 8851  | CCCAGACCCG | AGTTCACGGC  | CTAGTGGGAG | GGACAGGAAT | CTAGGCAGGC  |
| 8901  | AGATAATACA | GCGTGGTGCC  | TGCCAAGGCT | GGGGAGCCTA | GAGGCTGTAG  |
| 8951  | GAGTGCCGGG | GGGCTGGGGA  | AGTCTCCCTG | AAGAGGCTAC | TTATGATTCTG |
| 9001  | GGTCTGAGG  | GATGAGTAGA  | CTTCCCTGCT | CAGGTTTTGA | GGGATGGGCG  |
| 9051  | TGGAAACGA  | TGTGCCTGGC  | ATAGGCGTGT | ACTCTGAGTC | TGGGGAGAAG  |
| 9101  | TGGAGTCTGG | CTGAAGCCTC  | CAGTGGGCAG | AGGAGGGCCG | TGGTTAGTGA  |
| 9151  | AAGATGATGC | TGGAAACACT  | GTCCGGGCCA | CAGCATGAGG | GCTGGGAATC  |
| 9201  | CCTCCCTGA  | GGTCTTTGCT  | GACTGCATCC | TGCCAGCTCT | GTGAGGCCCT  |
| 9251  | GAGAGCTTTA | AGCATGGGGA  | GGGGCGTGAT | GGGATTTGTG | CCTGAGAAAG  |
| 9301  | CTCTGTCTGG | CAGCTGTGTG  | GTGGCTGGAT | TGGAGTGTGT | CATCGGAGGG  |
| 9351  | TGAGAGGCAG | CCAGCTGGCC  | AGGGAGGAGG | CTGTTTCTGC | AGCCCAAGTG  |
| 9401  | ACAGATGGTG | AGGCCTGGAT  | TAAGGCAGTG | GCAGCAGGAT | GGGGATAGGA  |
| 9451  | AGGAGGTGGG | GTGGTCAGCA  | TGGAGTGAAT | TGCCGGTCTG | GGGAGAGGAG  |
| 9501  | AGCCCTAGA  | CACCTAGGGT  | CCTGGCGTGG | GTTGGGGACC | AGGGGAGATG  |
| 9551  | CCCATCTCTA | AAATCTTAGC  | TTGGGCCAGG | CGCAGGGGCT | CATGCCTGTA  |
| 9601  | ATCCCAGCAC | TTTGGGAGGC  | CGAGGTGGGT | AGATCACCTG | AGGTGAGGGG  |
| 9651  | TTTGAGACCA | GCCTGGCCAA  | CGTGGCAAAA | GCCTGTCTCT | ACTACAAATA  |
| 9701  | CAAAAATTAG | CCTTGTGTGG  | TGGTGGGCAC | CTGTAATCCC | AGCTACTCGG  |
| 9751  | GAGGCTGAGG | CAGGAGAATC  | GCTTGAACCT | GGGAGGTGGA | GGTTGCAGTG  |
| 9801  | AGCCGAGATC | ACGCCATTGC  | ACTCCAGCCT | GGGTGACAAG | AGTGAAACTC  |
| 9851  | CATCTCAAAA | TAAATAAATA  | AATAAATGCA | TACATACATA | TATACATACA  |
| 9901  | TACATAAAAA | TAAAAAATAA  | AATCTTAGCT | TGGTTTCTTG | GGAGCATATT  |
| 9951  | CTTTCCTTGG | GGGAACAGGG  | TGGGGATCTG | GCTGAGGTTT | GACCTGCAGT  |
| 10001 | GACAGAAACA | GGACTGTCTT  | TATCCTGCTC | GAGCCTCTCC | TTTGCCTTCA  |
| 10051 | GATTAAGACT | CTCTTTGCAC  | ATATGGGGAA | ACTGAGGCAC | ACAGAGGGGA  |
| 10101 | GGGCTTTGCA | GAAAATCCCT  | ACCAAGGGCC | TAGAGGCATG | GGATGGGAAG  |
| 10151 | GGGACATTTT | ACCCCGGTAC  | GGTCAGTGGC | AGGCACAGTC | CTGTACCAGC  |
| 10201 | TTGGCTCCAC | CTCCTTTCTG  | TTGTAGTCCC | TTCTTTCCCC | TGAAGTCTTG  |
| 10251 | TTGCTGCTA  | TCCCCTAGCC  | TCCACAAAGA | AACGAGTTTA | TCTTACCTGG  |
| 10301 | TTCTTGGGTA | AAGCCTCATC  | AGGACCCAGC | TAATCACAGT | GAAGGGCTTC  |
| 10351 | CCTGGGGCAG | AACGGTTAGC  | GCCAGGGGCT | GGACAGGTGG | ATGAACAGAG  |
| 10401 | GCACGAGGGC | GCTGAAGACC  | TGCCTTGTGA | TTCTGGCCCC | AAGAAGAGAG  |
| 10451 | AGTTGAGGCT | GCCATGAGAG  | GGCTCGGTGG | TCAGGGCGGC | CCAGGCCTGG  |
| 10501 | TTCTCAGTTG | ATGGGGGCAG  | GTGCAACGAT | GCAGATGATG | AGAAGCAGTT  |
| 10551 | GGATCTGGAA | TAGATGTGAG  | AAGCTGAGCT | CACAGACCTT | GCTGATGAGC  |
| 10601 | AGGATGTGGG | GTCTCAGAGG  | AGGAATTGAG | GATGATCCTG | AAGTTTTTTGG |
| 10651 | CCTTTACACG | AATGGAAAAG  | AATGGGGAGC | AGCAGGGGCG | TTTTGTTTTG  |
| 10701 | CTTTGTTTTG | ATTTTGTGGG  | TGGTAGGCAT | TGCAGGCAGA | GAAATCAAGT  |
| 10751 | TCTGAATTAG | ACATGTTATT  | GCACTGTGTT | CAGATATACA | GAGACATATA  |
| 10801 | TCCATGCCAT | GCTGCCATAGT | TATCTACCAA | GATGCTATT  | GGAAATCTAT  |
| 10851 | GTGGGTAAAG | AGCTGGAGTT  | CAAGGGAGAG | GCTAGGGTTT | GAGATAAGAA  |
| 10901 | CATGAGACCA | CTTTCCATGG  | TCAAATGTCC | ACCCCTCTGA | GCTTCTGTGC  |
| 10951 | CCTGAAGGGT | GTGTCAGATT  | CCTTGTGTGT | GCCTGGCACA | TAGTAGGCAA  |
| 11001 | TCAAGAAAGT | GCCACTGGTT  | TTATGGTTAT | TGTTATACGG | CACCCGCCTT  |
| 11051 | CTCTGCCCGC | AGCCTCCCTC  | TCCTCTTCTC | CCTTCTCTCT | TCTTCTCTCG  |
| 11101 | CCTTCTCTCC | TCCCTCCTCT  | CCAGCATCCT | GGGGTCCGTT | GGTCCAGATG  |
| 11151 | AAGGTAATTG | CCAAGGAGGG  | AGCCACAGG  | TCGATGGTCG | CGGGATGGGG  |
| 11201 | TCAGTGGGGT | CATTGTCTCT  | CTTGGCTGGG | ACCTTACCAG | TCATGTCAGC  |
| 11251 | TTGAGCCACC | TGTCACCTCG  | TGGTGGTGCT | GGGGCCAGAA | AGCAGGGCAG  |
| 11301 | ACCTCCAGCC | TATTAGGTCA  | TTTCTGATTT | GGGATTCGTC | CTACTATATG  |
| 11351 | TGGCTGACCT | TACACCCAG   | CTGTGTCATC | CTGCTTGTC  | CAAGGCCTGG  |
| 11401 | GGTGCCATCC | ATCTCTCTGA  | AACCCCATCA | GCCCAGATCC | CGAGGGCTGA  |
| 11451 | GATGGTACCT | CTGTAGGATA  | GCAGAGTCCC | TACAATCTTA | CTCTCAGTCC  |
| 11501 | CAGCAGCAGG | GACATCTTTG  | CCTAGCCTGG | GTGGGGGATG | GAACTGGAGA  |
| 11551 | AAGGTTTTGA | TTGGCTTTGG  | GCCTGCAGAC | GGCACTCACA | GGGAAGGGGC  |
| 11601 | AGAGCTAGCC | TAGGAAGAAC  | TCTGCTCCCA | GCTGGGGGCG | GTGGCTCACG  |
| 11651 | CCTGTAATCC | CAGCACTTTG  | GGAGGCCGAG | GTGGGTGGAT | CACCTGAGGT  |
| 11701 | CAGGAGTTCA | AGACCAGCCT  | GACCAACATG | GCGAAACCCT | GTCTCTACTA  |
| 11751 | AAAATACAAA | AAGTAGCCGG  | GCGTGGTGCC | AGACACCTGT | AATCCCAACT  |

FIGURE 3D

|       |            |             |             |            |            |
|-------|------------|-------------|-------------|------------|------------|
| 11801 | ACTCGGGAGG | CTGAGGCAGG  | AGAATCTCTT  | GAACCTGGGA | GGTGGAGGCT |
| 11851 | GCAGTGAGCC | GAGATCACGC  | CATTGCACTC  | CAGCCTGGGG | GACAGAGTGA |
| 11901 | GACTCTGTCT | CAAAAAAAAA  | AAAAAAAAAAC | CAAAAAAAAA | AACAGCAACA |
| 11951 | ACTCTCCTGC | CCTAGTTTCC  | TCTGACCTCC  | CCACTCAGCA | GCAGATCCCT |
| 12001 | TGTTTGTGTC | GGAGAGGGTG  | CTGGACTTGG  | AGTCCAAAAG | CTCCTAAGAT |
| 12051 | TCCAGTCCTG | GCTCTGCTGC  | TCACAGCCTG  | GGCTCAGTGT | CTGCACCTGC |
| 12101 | GTGGAGCAGA | TGGCCCTGAC  | GTCCTCCTCC  | CAGGTCGTCA | CCAGACGAAA |
| 12151 | GTGTGCATGG | GCTGGGATGT  | CCCGGCCCGC  | GTCCCTGGCT | GTGCAAGGAC |
| 12201 | GGGTGTGGGG | TCCTGGCCAG  | CGGTGCCCAG  | GCCAGCGCTC | AGCTCAAGCT |
| 12251 | CCCCTTCTCT | GCAGTGCACT  | GCCAGAAGAT  | CGTCCACAGG | GACATCAAGC |
| 12301 | CATCCAACTT | GCTCCTGGGG  | GATGATGGGC  | ACGTGAAGAT | CGCCGACTTT |
| 12351 | GGCGTCAGCA | ACCAATTTGA  | GGGGAACGAC  | GCTCAGCTGT | CCAGCACGGC |
| 12401 | GGGAACCCCA | GCATTTCATG  | CCCCCGAGGC  | CATTTCTGAT | TCCGGCCAGA |
| 12451 | GCTTCAGTGG | GAAGGTGACT  | CGCAGGCCCT  | GGGCCAGGCT | GGGGTTCAAG |
| 12501 | TGGGGGGCGT | AATAGCTTGC  | CGCAGTGGCC  | CAGTTTCTAA | CCTGAGGGTG |
| 12551 | CCAGGTCTTT | TGTGTCTAGG  | GAGTGACATA  | TTTGCCTCTT | CCTTGGAGCC |
| 12601 | TGACAAACTC | CACAACTTTG  | GCCTTCTCCT  | GTTTTCCAGC | AAAGTGGTCC |
| 12651 | CAAATCTCCC | TTGCAGATAT  | TTACTGTTGG  | TTGCTCTGTG | CTGGGTTCTG |
| 12701 | GACCCGACTG | TGGAAGAGGC  | AGAAACAAAG  | AGAACCCTGT | TTCTGCCCTT |
| 12751 | CTGGATGGTT | TCGGGGGAAG  | TTGGGGGTCC  | CCGCAGATCT | TGGGACATGG |
| 12801 | CAGGATTTGA | ACTGGCCCTT  | GAAGAATGGG  | GAGGATCTGA | GCAGGACCTG |
| 12851 | GAGCCTAGAG | AATAAACCCAG | AGAACAGAAG  | GGCTCAGGGT | GGGGGGCAGA |
| 12901 | GGGTATAAAG | GGCCTGGAAG  | TTTGGGCTTT  | CTCCTAAGTG | ACAGGAGCGT |
| 12951 | AGGCAAAAGT | GTCTGAACAA  | GAGGTTACAC  | GGTCTGGCGC | AGTTCCCTGG |
| 13001 | GCACATGGCT | GTTTCACCTA  | TGGAGTGCCA  | GCCACCCAC  | TGCCAGGGAG |
| 13051 | GCTGTGGGTG | AGAGGCATTT  | GGACACGTGT  | GAGTATCCAG | GAAAGAGGTC |
| 13101 | AGGAGGCCGG | GCACAGTGGC  | TCATGCCTGT  | AATCCCAAGT | CTTTGGGAGG |
| 13151 | CCAAGGTGGA | TCTCTTAAGG  | CTAGGAATTT  | GAGATGAGCC | TGGGCAACAT |
| 13201 | AGCAAGACCC | CATTTCTACA  | AAAAAAAAAAA | TAAAAACATT | AGACAGGTGT |
| 13251 | GGTAGTGAC  | ACCTGTAGTC  | CCAGCTACTT  | GGGAGGCCGA | GGTGGGAGGA |
| 13301 | TCGCTTGAGT | CCAGGAGTTG  | GGGGCTGTAG  | TGAGCTGTGA | TGGTGTCTAG |
| 13351 | CCTGAGTGAC | TGAGCGACAC  | CTTGTCTCGA  | AGAAAGAAAG | AAAGACGTTG |
| 13401 | GGGATGTTGA | TAAAGATTTT  | TTGAAATGTT  | TTATTTTGAT | ATAATTCTAA |
| 13451 | ATTTACAGAA | AAGTTGGAAG  | AATAGTACAA  | AGAAATCCCC | TATATCTTTT |
| 13501 | TACCCAGATT | CACCAATTAT  | TGACATTTTG  | TCCCACTGGC | TTTTTCATCA |
| 13551 | TCTTTCTTTT | TTTTTGAGCC  | GGAGTCTCGC  | TCCTGTCCGC | CAGGCTGGAG |
| 13601 | TGCAGTGGCG | CGATCTCAGC  | TCAGTGCAAG  | CTCCACCTCC | TGGGTTACAG |
| 13651 | CCATTCTCCT | GCCTCAACCT  | CCCGAGTAGC  | TGGGACTACA | GGCGCCACCC |
| 13701 | ACCACGCCCG | GCTAATTTTT  | TGTATTTTTT  | AGTAGAGACG | GGGTTTCACC |
| 13751 | GTGTTAGCCA | GATAGGTCTG  | GATCTCCTGA  | CCTCGTGATC | CGCCCGCCTC |
| 13801 | GGCCTCCCAA | AGTGCTGGGA  | TTACAGGTGT  | GAGCCACCAC | GCCCAGCCAG |
| 13851 | AAATTTATCA | TTGATAAGAC  | TTATATATCG  | GTCAGGCATG | GTGGCTCATG |
| 13901 | CCTGTAATTC | CAGCCCTTTG  | GGAGGCCAAG  | GTAGGTGGAT | CACCTGAGGT |
| 13951 | CAGGAGTTTG | AGACCAGCCT  | GGCCAACGTG  | GTGAAACCCC | GTCTCTACTA |
| 14001 | AAAAATACAA | AAATTAGCCG  | GGCATGGTGG  | CGGGCACCTG | TAATTCCAGC |
| 14051 | TACTTGGGAG | GCCGAGGCAG  | GTGGATCACC  | TGAGGTCAGG | AGTTTGAGAC |
| 14101 | CAGCCTGGCC | AACGTGGTGA  | AACCCCGTCT  | CTACTAAAAA | ATACAAAAAT |
| 14151 | TAGCCGGGCA | TGGTGGCAGG  | CACCTGTAAT  | TCCAGCTACT | TGGGAGGCTG |
| 14201 | AGGCAGAAGA | ATCGTTCGAA  | CCCAGGAGGC  | AGAGGTTGCA | GTGAGCTAAG |
| 14251 | ATCGTGCTAT | TGCACTCTAG  | CCTGGGCGAC  | AGAGTGAGAC | TCTGTCTGAA |
| 14301 | AAAAAAAGA  | CATACATAAT  | CCACAGACCT  | TATTTAAATG | TTATCAGTTG |
| 14351 | TCCTGATACT | GTAATTCATA  | ACTTCTTCTT  | TTTCTGGTCC | CGGAATCCAA |
| 14401 | TCGAGGACCA | CTTGCTGCAT  | TCACCTTCTT  | GTCTGTGGTA | TCCTTTCATC |
| 14451 | TGGAAGAGGG | CCTTGGCCTG  | CCGTTGTCTT  | TCCTGATCTT | GACATTTTGG |
| 14501 | AAGACAACCA | GCCTGTTATT  | TTGTAGAATG  | TTGTCAGTTT | GCATTTGTCT |
| 14551 | GGTGTTCCTT | GGTTGGGATT  | CAGATGATGC  | ATCTGGGGCA | GGAATATGTA |
| 14601 | GGTAGAGATC | GAGAATCACT  | CATATAAGCG  | AGAAAGTGGA | TACCAGAAGA |
| 14651 | GGTGGCGTTC | CGGAGCAGAA  | GGTAGAGAGA  | GCACACGCTG | GAGTCCAGGG |
| 14701 | CGCGGGGAGG | CCCAGGGGTG  | TTTGGGAGCC  | CAGAGGAGTT | GTTGCAGTGG |

FIGURE 3E

|       |             |             |            |             |            |
|-------|-------------|-------------|------------|-------------|------------|
| 14751 | CGGTGGATGA  | GGGCGTGAGA  | GGACAGGGCC | TCTGTGTGGG  | CAGGGGCTGT |
| 14801 | TTGCAATATC  | AGGAAGAAGG  | TGGATTATGA | GGAGAAGGGA  | TGACTCCTTG |
| 14851 | AAGCCCAGGC  | TGGTTTAGTG  | AGCAGAAAGT | CCATATATAC  | CATCATTCTT |
| 14901 | GGGGTGCCTG  | TGTGGCACGG  | GAGCGGCCCG | TGTGACCCTC  | TGGATGAAGG |
| 14951 | AGGTTTTGTA  | CCTGTTGAGT  | TGGAAACGTA | CCTGGTTAGA  | GTCTTTCCCA |
| 15001 | AGGAAACCCA  | GAACCCCTGG  | AGGGTGGAGG | CCTTGTTCTG  | GCCGCCCTTG |
| 15051 | TGTCCTCAGC  | ACTCAGCACG  | GGGCCCAGCA | TCGGGCAAGT  | ACCGCGGAGT |
| 15101 | GTTTGTCTGAG | TGAGTACATG  | ACAGAGGAAA | GAGGTTCCCT  | GCAGGCCTCT |
| 15151 | CCTGCAGCCC  | GCTGGAGCTG  | GGTGGGCAGA | GGTGGCTGTG  | CCTGTTGGGG |
| 15201 | ACTGATGTGA  | GCATGTTTCT  | TTCCAGGCCT | TGGATGTATG  | GGCCACTGGC |
| 15251 | GTACAGTTGT  | ACTGCTTTGT  | CTATGGGAAG | GTGAGTGCCA  | GGGATGCCAG |
| 15301 | CAGAGCTGGG  | GCGGGTCCAG  | TGAGGCGGGC | ACGGGCGACG  | GATGCAGGCT |
| 15351 | CTTCCTTTTT  | GTCCTTAAGT  | GGCTTTTGAA | AGAGCCCACC  | TGGCTCAGAG |
| 15401 | AAGGCTGAGA  | GAGAAGAGGC  | TTTTTCTATC | TTTCTCTGGT  | CCCCTGCGGA |
| 15451 | GCGATTCTCG  | CGAAGGAGTC  | GCAGGACAGC | AGACACCTAA  | GGGGAGGTGC |
| 15501 | CGACGATGGT  | GTTGCCACCG  | CCCCAGCCAG | AGTGCTCCCC  | GTCCCTCTGT |
| 15551 | CCCTTGACCG  | CATTCACTTA  | TTGAGCCATG | TGTTCACTCC  | CTTGCTCATT |
| 15601 | TATTCGACAA  | ATTGTCTTTC  | ACCCCTACCC | TGGCTGAGGC  | TGGACCCTGG |
| 15651 | GGACACCCAA  | CGCTGACGTA  | TCGGTGATCC | CTGCCCAGAG  | GTGTGCCTGC |
| 15701 | TCTGGTGACC  | ACACTAAGGG  | GCAGGGGGGA | ATTTCAAGTGA | ACATGTTCCC |
| 15751 | AAGCCCCAGG  | CCCTGGGAGT  | GGAGGCCTGG | CCACAGGTGG  | CGGTAATGGT |
| 15801 | GGTGGGTGCA  | CCCAGCCTGG  | CCTGGCTTGG | CCGCGGGTGG  | CAGTAACGGC |
| 15851 | GGTGGATGCA  | CCCAGCCTCA  | TTGTTCCCTC | AGCAACTCAT  | TCATTAGTTC |
| 15901 | AACATTTGTT  | GAACATTTAC  | AGTGTGAGTT | GAGGTCCTTC  | TCATGTAATG |
| 15951 | GGAGCCCAGA  | CCTGCCCCCT  | ACCCCTGCCC | CCACCAAGGG  | AGGGGGGTTG |
| 16001 | ATCCCCTGGC  | ACAGGTCGAG  | GCCCTGGACC | CACATCCTTT  | GTCTGCCTCT |
| 16051 | CCACCCACAA  | GTGCCCCTTC  | ATCGACGATT | TCATCCTGGC  | CCTCCACAGG |
| 16101 | AAGATCAAGA  | ATGAGCCCGT  | GGTGTTCCTT | GAGGGGTGAG  | TTGTCCACCC |
| 16151 | AGGGGAACAA  | GGGGGCTACC  | ACCCGCTCCT | GGTGTCTGAG  | TTTTAGCAGA |
| 16201 | GCTTTTGCCC  | TCTGAGGACC  | CCACCCAGC  | CTGCAGATAT  | GAAGGTGGCG |
| 16251 | GTGCTGTTCC  | CTGGGAGGGA  | CCCCTGAATA | GATGGACGGG  | AGGGACTCTG |
| 16301 | GAGCCAAGGG  | TCTCCGCAAC  | GTCAGTGTGT | GGATGGGAAC  | CCTGAGATCC |
| 16351 | AGGGTTGGCC  | AGGGATGACC  | ACAGGCATCA | TTACACCCAC  | TCCTTCACCG |
| 16401 | CAGGCCTGCC  | TGGGGTCAAGT | GGCGCCAGCC | CCACCCAGCC  | CCTGGACTCA |
| 16451 | AGGGGAACCT  | CTCCTTCCCC  | CATCAGGGT  | CAGGGAACTT  | CAAGATGCCA |
| 16501 | GTGCGTGCTC  | CCCATTTTAC  | AGATGGAAAA | GAGGATGCTC  | TGGAGGAGAG |
| 16551 | CGGTCAGGGG  | GCTGGGACTC  | AAGCCACTCT | TCCTCCCCAC  | TCTTCCCATT |
| 16601 | GTGACCGAGC  | TCTCTGAGCG  | TAGCAGGGAT | GTGCGGGAGG  | CCTCTTGCTC |
| 16651 | ATGCATGGTT  | GCCTCATGTA  | CGGCCACCGT | GGCAGCCACA  | GCCTGAGCTC |
| 16701 | CCAGGCTCCT  | CTTTTCAGCA  | GTGGATTTCA | GGAGTGAAAT  | GGAGGCCGGG |
| 16751 | TGCGGTGGCT  | CACGCCTGTA  | ATCCCAGCAC | TTTGGGAGGC  | TGAGGTGGGC |
| 16801 | AGATCACCTG  | AGTTAGGAGT  | TAGAGACCAG | CCTGGCCAAC  | ATGGTGAAAC |
| 16851 | CCCATCTCTA  | CTAAAAATAC  | AAAAATTAGC | CAGGCGTGTT  | GGCGCACATC |
| 16901 | TGTAGTCCCA  | GCTACTCGGG  | AGGCTGAAGC | ACGAGAATTG  | CTTGAACCCA |
| 16951 | GGAGGCAGAG  | GTTGCAAGTA  | GCCTGGGCGA | CAGAGCAAGA  | CTCTGTCTCA |
| 17001 | AAAAAAAAAA  | ACAGAAGAAA  | GAAACTGAAT | AAGGCCGGGT  | GCGGTGGCTC |
| 17051 | ACACCTGTAA  | TTCCAGCACT  | TTGGGAGGCC | AAGGAGGGCG  | GATCACGAGG |
| 17101 | TCAGGAGATC  | GAGACCATCC  | TGGCTAACAC | GGTGAAACCC  | CATCTCTACT |
| 17151 | AAAAATAGAA  | AAAAAATTAG  | CCGGGCGTGG | TGGCGGGCGC  | CTGTAGTCCC |
| 17201 | AGCTACTCGA  | GAGGCTGAGG  | CAGGAGAACG | GCGTGAAACC  | GGGAGGCAGA |
| 17251 | GGTTTCAGTG  | AGCTGAGACC  | GTGCTACTGC | ACTCCAGCCT  | GGGCGACAGA |
| 17301 | GCGAGACTCT  | GTCTCAAAAA  | AAAAAAAAAA | AAAAAAAAAC  | AAAAAAAAAA |
| 17351 | AACAAAAAAC  | AACAAACAAA  | AAAAGAAAAT | GAAACGGGAC  | TTGTACTCAG |
| 17401 | CGACTCCTGC  | TCTCTTCTGC  | TTATTTCTTG | TGTGGTCCCC  | AAGCCCTGCT |
| 17451 | GAGCCCTCCT  | CTTCCCTGTC  | TCTGGGCCCT | GTTGCCACTT  | ATACCCCTTG |
| 17501 | CCTCATTCAG  | GCCTCAGGCC  | CCTCCCCAGA | CTTATCTAGC  | CACCTTCCCC |
| 17551 | CTGGTCTCGC  | TGCTGCTGGC  | CTCCCTCCAG | TCCAGCCAAC  | ACATTGAGGC |
| 17601 | GGGGACAGCC  | CTGATAAAGC  | ACAACAAATC | TGCCTGCATC  | TCTTGCTGTA |
| 17651 | AGTTTGTCTG  | AAGCTTCTCA  | AAGCCACACC | CTGGCGCTAG  | CATTACACGC |

FIGURE 3F

|       |            |             |             |            |            |
|-------|------------|-------------|-------------|------------|------------|
| 17701 | TCTCCGGGTT | CTGCCACCCG  | CTCGTCTGGG  | GCCGCCTCAC | TCCCTTTCCC |
| 17751 | GAGCACCAGC | CAGCTGGCTT  | CTGTCCATTT  | CCTCCTCATC | CTGTGGTTGC |
| 17801 | CTTCCCTCCC | TGCCTCCACA  | GTTGTACCCC  | TGGTGCCTCT | CTTCCTGCTA |
| 17851 | TACCCCTGTC | TGAGGGGTGT  | CTTTCCCCTC  | AGCCCAGGAA | TTTTAAAAGG |
| 17901 | GATGAAGCAT | CTAAGACAAC  | AGGGGGAACC  | GAAGTCAACA | GTCTTGAGAG |
| 17951 | TGGCTTTCTG | CTCCCTACTC  | TTGGAAGGAT  | GGGCTCCCCA | AGACCACTGG |
| 18001 | TGGCAAAGAA | ACCTGGGGTT  | TGGCCGGGCG  | TGGTGGCTCA | CGCCTGTAAT |
| 18051 | CCCAGCACTT | TGGGAGGCCA  | AGGCAGGCGG  | ATCATGAGAT | CAGGAGATCG |
| 18101 | AGATCATCCT | GGCTAACACG  | GTGAAACCCC  | GTCTCTACTA | AAAATACAAA |
| 18151 | AAATTAGCCG | GGCACGGTGG  | CGGGCACCTG  | TAGTCCCAGC | TACTCGGGAG |
| 18201 | GCTGAGGCAG | GAGAATGGCA  | TGAACCTGGG  | AGGCGGAGCT | TGCAGTGAGC |
| 18251 | CGAGATTGCG | CCACTGCACT  | CCAGCCTGGG  | CCACAGAGCG | AGACTCCATC |
| 18301 | TCAAAAAAAA | AAAAGAACCC  | TGGGGTTTTGG | GCAGAGAGAG | TTGGAGCTGA |
| 18351 | TGTGGCGCTG | AGGGGGCTGC  | TCCCTCCCAT  | CTGAGTCTCC | CATCTCTGCC |
| 18401 | TGCACTCTTC | TGGCTGGCAC  | TGTGCCAGCC  | TGCTAACCTC | CCTGGGCCTC |
| 18451 | AGTTTCCTCC | TCTGTCAAAT  | GAGAGAGGAT  | CTTCTCTGGG | TGTAGAAAAG |
| 18501 | GACGAGGTGG | TGAGTGGGTC  | TGAAGGCCTC  | TGGTGTCCCA | TAAAGCGACT |
| 18551 | CTCCTCACCA | TCTTTGCCAC  | CCATTGGGGT  | GTCCAGCAC  | CATGGAACTC |
| 18601 | TGTCTGTGCC | TCTGTCTCTG  | AGGGAGACTT  | GACCTCCTGC | TCAGGAAAGG |
| 18651 | CTCTCCAAGC | CCTTGTGTGT  | AAATTCCTGC  | CTGCTGTCCG | GAACCTCAGT |
| 18701 | TTCCCATCCG | AGGGACGAAG  | GTTTCGGGAA  | GAGAGGTGGA | CAGGAAGGGG |
| 18751 | TCCTCATCAG | CGGTCCCACC  | CTCCTCTCCT  | TCCTTCGCCC | TCTCCAGGCC |
| 18801 | AGAAATCAGC | GAGGAGCTCA  | AGGACCTGAT  | CCTGAAGATG | TTAGACAAGA |
| 18851 | ATCCCGAGAC | GAGAATTGGG  | GTGCCAGACA  | TCAAGGTCGG | GGAAGTGGGG |
| 18901 | GTCTTGGGCT | GGGCTGGGAC  | ACAGAAAAACA | GGAGTCACTT | TCCCTTTCTG |
| 18951 | GAGGGATCAA | CACCAGGATG  | CATGTGTGTT  | GGGTTTGAGT | CTGTGGACTT |
| 19001 | TGGACCCCTC | CAGGTGATTG  | TGGTAATGGC  | CTGACCTCTC | CCCCTCTCCC |
| 19051 | TGCCCTCCCC | GCCCCGACAG  | TTGCACCCTT  | GGGTGACCAA | GAACGGGGAG |
| 19101 | GAGCCCCCTC | CTTCGGAGGA  | GGAGCACTGC  | AGCGTGGTGG | AGGTGACAGA |
| 19151 | GGAGGAGGTT | AAGAAGCTCAG | TCAGGCTCAT  | CCCCAGCTGG | ACCACGGTGG |
| 19201 | TAAGAGAGCC | GGGGTAGATG  | CTCCCTTGTC  | CTGGAGGGCC | TGGGGGACCT |
| 19251 | GAGCCTTGCT | CTGTGCCTGG  | CTCCTTGGGG  | GGACAGAGGC | CTGCCTGGCC |
| 19301 | AGCCAGCTGT | GATCCTGGGC  | CACTGGAGCC  | GCCATTCTGC | TGGAGGCCCA |
| 19351 | TGGAGAGGGA | GGTCTTGTGG  | TCGGGAGACC  | AGGAGGCTTG | GTGAGGAGAG |
| 19401 | TGACTGATTT | AAAGAAATAG  | CGGGCGTGGG  | GCCGGGCGCG | GTGGCTCACG |
| 19451 | CCTGTAATCC | CAGCACTTTG  | GGAGGCCAAG  | GCGGGCAGAT | CACGAGGTCA |
| 19501 | GGAGATCGAG | ACCATCCTTG  | AAACCCCGAC  | TCTACTAAAA | ATATAGAAAA |
| 19551 | TTAGTGGGCG | GTGGTGGCGG  | GCGCGTGTAG  | TCCCAGCTAC | TCGGGAGGCT |
| 19601 | GAGGCAGGAG | AATGGTGTGA  | ACCCGGGAGG  | TGGAGTTTGC | CGTGAGCCGA |
| 19651 | GATCCGGCCA | CTGCACTCCA  | GCCTGGGCCA  | CAGAGCGAGA | CTGCGTCTCA |
| 19701 | AAAAAAAAAA | AAGAAGAAAA  | GAAAAGAAAG  | AAATACCGGG | CGCGGTGGCT |
| 19751 | CACGCCTGGA | ATCCCAGCAC  | TTTGGGAGGC  | CGAGGCGGGT | GGATCACGAG |
| 19801 | GTCAGGAGAT | CGAGACCATC  | CTGGCTAATA  | CGGCGAAACC | CCACCTCTAC |
| 19851 | TAAAAATACA | AAAAAATTAG  | CCGGGCGCAG  | TGGTGGGCAC | CTGTAGTCCC |
| 19901 | AGCTACTGGG | GAGGCCGAGG  | CAGGAGAATC  | GCTTGAACCT | GGGAGGTGGA |
| 19951 | GGTTGTAGTG | AGCCAAGATC  | ACGCCATTGC  | ACTCCAGCCT | GGTTGACAGA |
| 20001 | ACGAGACTCC | ATCTCAAAAA  | AAAAAAGAAA  | GAAATAGATG | GCCCTTGCTC |
| 20051 | AGCGGCAGCA | GTCACCGTGA  | CTGGAAGAAG  | CATTTTCATT | CGTCCAGACA |
| 20101 | GTTACTGAGC | TTCCGTTCTC  | CAGGCACTGC  | ACAAGGTGCC | GAGGACAAGG |
| 20151 | CAGGGGAACG | GCCTGGGCAG  | CCTTTGGATT  | GGAGGAGTGG | CCCCAAAGCC |
| 20201 | CACGTATCAG | TTAGGCGGCG  | CCTGCGTCTC  | CCCCAGAGCC | CACGTATCAG |
| 20251 | TTAGGCAGCA | CCTGCGTCTC  | CCCCAGAGCC  | CACATATCAG | TTAGACGGCG |
| 20301 | CCTGCTTCTC | CCCCAGCGCC  | CACGTATCAG  | TTAGACGGCG | CCTGCTTCTC |
| 20351 | CCCCAGAGCC | CACGTATCAG  | TTAGACGGCG  | CCTGCTTCTC | CCCCAGATCC |
| 20401 | TGTGTATCAG | TTAGACTGCG  | CCTGCTTCTC  | CCCCAGAGCC | CACGTATCAG |
| 20451 | TTAGACGGCG | CCTGTTACTC  | CCCCAGAGCC  | CACGTATCAG | TTAGACGGCG |
| 20501 | CTTGCTTCTC | CCCCAGATCC  | CGCGTATCAG  | TTAGACGGGC | CTGCGTCTCC |
| 20551 | CCCAGATCCC | GCGTATCAGT  | TAGACGGGCC  | TGCGTCTCCC | CCAGAGCCCA |
| 20601 | CGTATCAGTT | AGACGGGCCT  | GCGTCTCCCC  | CAGAGCCAC  | GTATCAGTTA |

FIGURE 3G

|       |            |             |             |            |             |
|-------|------------|-------------|-------------|------------|-------------|
| 20651 | GACGGCGCCT | GCTTCTCCCC  | CAGAGCCCAC  | GTATCAGTTA | GACGGGCCTG  |
| 20701 | CGTCTCCCCC | AGAGCCCACG  | TATCAGTTAG  | ACGGCGCCTG | CTTCTCCCCC  |
| 20751 | AGAGCCCAGG | TATCAGTTAG  | ACGGTGCCCTG | CATCTCCCCC | GTGCCCACGT  |
| 20801 | ATCAGTTAGA | CGGCGCCTGC  | TTCTCCCCCA  | GAGCCCACGT | ATCAGTTAGA  |
| 20851 | CGGGCCTGCG | TCTCCCCCAG  | ATCCTGCGTA  | TCCATTAGAC | AGTGCCTGTG  |
| 20901 | TCTCCCTAG  | TGCCCGCTCA  | CATTTTCGGT  | TTGCTCCTCT | TCCTCTGCTC  |
| 20951 | AGCTTCTGTG | TTGGCACTTG  | GAAAGTGATT  | CACATAGTCC | CCCGTGGCCA  |
| 21001 | CCTGGGGCCA | CTGAGAGCCC  | TGCCCTGCCC  | CTGCCTGACA | GTCAAGTGAG  |
| 21051 | TCAGGGCAAG | CACAAGGCCA  | GGAGGAGAGC  | CAGGGCCACT | GCCGTTGGCG  |
| 21101 | GGGCTGGGCC | TTGCACTTTA  | TCCCCCTCTG  | CAGGGTCCCG | GCCCAGCTGG  |
| 21151 | GACCAGCTGG | CTCAATCCCT  | GCCCCCTATG  | CTTACTTGAC | TCTGTGGGGT  |
| 21201 | CGCTGGAACC | AGGCAACTCC  | CACGGGGTCC  | CCATGACCAC | TTGCCTGATC  |
| 21251 | TTAGCCACCA | TCTCCTCTCT  | CTCAGACCAC  | TGGAACAACC | TCCCACGCTG  |
| 21301 | TCCCTTGCTT | CTACTCTCAC  | TCCCTGTCCC  | CCTGGTCAAT | GCTCAACTCA  |
| 21351 | GCACCCAGCA | TGGTCCCAGT  | GGCATGAGTG  | TGTCACCTCC | CAGCTCAGAG  |
| 21401 | CCTGTCTCTC | ACTCGGGCTG  | CTGTGTCCCT  | CAGAATCAGA | CCTCCAGCCT  |
| 21451 | GTGCCCCACC | ACCCGCCCTG  | TTTTTCTGCG  | GGGCTCGTGC | ACCGTCCCCG  |
| 21501 | CATCATGCAC | TCGTCTCTGG  | CCACGTGCCA  | TGGAAGGGGC | TGCCCCAGAG  |
| 21551 | CCTTCAGACT | TCGCTTCCCT  | CTGCCCCGGG  | AGTCCCAACC | CCGATGGCCA  |
| 21601 | CGGGACTCGC | TCCCTCACTT  | CCTTCGGCTT  | TTTACGCCAG | GGTCCCCTCC  |
| 21651 | TAGAGAGAAG | CGAGCCTTCC  | CTGACCCTGT  | AGCTTCAGCC | TCCCCTGCTT  |
| 21701 | CACACCTCAT | CGCCATTCCC  | TTGTTTTATT  | TTTTCTTTTC | CACTTACTGA  |
| 21751 | CATACATAAT | TTACTGATTT  | TTCTTCTTTA  | CTTATCGCCT | GTCTCCCCCA  |
| 21801 | ACTAGAATAT | AAGCTGTATG  | ATGGCTGGGC  | GCAGTGGCTC | ACGCCTGTAA  |
| 21851 | TCCCAGCACT | TTGGGAGGCC  | AAGGCGGGAG  | GATCACTTGA | GGTCAGGAGT  |
| 21901 | TTGAGACCAG | CCTGGCCAAAC | ATGCTGAAGC  | CCCGTCTCTT | CTAAAAATAC  |
| 21951 | AAAAAATTAG | CCGGGTGTGG  | TGGTGGACGC  | CTGTAATCCC | AGCTATTGAG  |
| 22001 | GAGACTGAGG | CGGAAGGATC  | ATTTGAAGCG  | GGGAGGCAGA | GGTTGCAGTG  |
| 22051 | AGCCGAGATT | GTGCCACTGC  | ACTCCAGCCC  | TGGGCAACAA | GAGCAAAACT  |
| 22101 | CCGTCTGAAA | AAAAAAAAGG  | CTATATGAGG  | GCAGGAATTC | TGGCCTCAGT  |
| 22151 | GTGGCCCCAG | GGCCTAGAGT  | AGTGGCCAGC  | ACCCAGTAGG | CAGCCAGTGG  |
| 22201 | TGACCAAGTG | TGACGGGATG  | GATGGACACA  | AGCGAGGGAG | TGAAGGGACT  |
| 22251 | GGCAAGTGTG | CCGCTGCCTC  | TCTGCATGCG  | TGTGAGTCGG | CGTGTCTGTG  |
| 22301 | GGCACGGCAT | GGAACCGTCC  | TTGTACCGGA  | GGAGGGACAA | AGGCAGAGAG  |
| 22351 | CCAGGCTGCG | GCAGCTGTTT  | CCCTCCTGGC  | AGCCCCACTG | ACTGGGCCAC  |
| 22401 | CGGCTGCGGC | TCAGCCGCTT  | CCCGGGCCGC  | CCTGCAGTAG | CATCTTGGCA  |
| 22451 | TCTTCTCGGC | GGCCGGAAGG  | CGGGAAGGAT  | GGCACAGCAT | CCCTCCATGG  |
| 22501 | CATTGCTGCC | GTACGAGAAA  | GGTATCTTCT  | AATGGACTCC | CACTTCCAGC  |
| 22551 | CCTGGCCCTC | CCCACTCTTT  | CAGCCTGGCC  | TTGCGGACCC | TTTATGGGCT  |
| 22601 | GGTCCCGGCC | CCCTCCTCAT  | GTACCAAGTG  | CATCCGGCTC | CTCACCATTG  |
| 22651 | CAGGAATATG | CCCCCAGCTG  | CCAGCGCCCC  | GTGTTCTTGC | CTCTGCCATT  |
| 22701 | TCATGCTGTG | CTGATTGAGA  | TGGGACCCGC  | ACTGCGGGCC | CCTTGGCAGC  |
| 22751 | TGCTCTCGGG | GAATCGGAGC  | AGAGGCTGCG  | TGTCTGGGAG | CCTGGGACCT  |
| 22801 | GTGCTCCTCA | CGCTGCCTTG  | TCCTCCTCAG  | ATCCTGGTGA | AGTCCATGCT  |
| 22851 | GAGGAAGCGT | TCCTTTGGGA  | ACCCGTTTGA  | GCCCCAAGCA | CGGAGGGAAG  |
| 22901 | AGCGATCCAT | GTCTGCTCCA  | GGAAACCTAC  | TGGTGTAAAT | ACTGGTGGGC  |
| 22951 | CAGGGACTGC | CGGGCACTCC  | CTGGAGTTGG  | GTGGGGAGGT | CTGAGGCCCA  |
| 23001 | TCCTCCCACT | CTCACTGTCT  | TTGGGCCAAG  | GCCAGAGCCT | GGGGACTTGG  |
| 23051 | CCAGTCTCG  | GTGTTGGCCC  | CATTTGCATC  | TCTGTCCCCA | AGGTTAGTCG  |
| 23101 | GGGCTAGAAG | GGACCTTTTG  | GGCCAGCTC   | TTGCTTCATT | CCTGGGGCCA  |
| 23151 | GCATCCCTCA | CACACACACT  | TCCAGGGATG  | AGGAGCTCAC | GCAGCCCCCTC |
| 23201 | CATGGGACAG | GAAGACCCTT  | CTTCCATGCA  | GCTTGATGTC | ACTCTCTCAC  |
| 23251 | TGGGTCCAGC | CCCTCTGGGG  | CTTCAAATCT  | GTGGCCCCCT | CAGCCCTTGG  |
| 23301 | CAGCCTGGCA | GAGGTTTGCA  | GACAGGCTGA  | TGTTGGCTTC | CTGTAGGAGG  |
| 23351 | CTGGCGGGCT | GTAGAGGAGG  | GGTGCTGGCC  | CCTCTGCCTG | GCCCTGGGGA  |
| 23401 | CTGTTGGCTG | CTCTCCCAAG  | TGGCCCAAGC  | TGCCTGCAGC | CATTGCTGGG  |
| 23451 | GCTCTGTGCC | CAGTCAGCAC  | TTTGTGAGTG  | CTTGTTCAGT | GAGTAAGCAG  |
| 23501 | GGACAGGCTG | GCCGGTGGAC  | CACGGGAGAG  | GAACCCGCAT | TGGCCGAGGG  |
| 23551 | CTCCCTATGG | TGAGCCACGC  | CTGTGGGTTC  | ACCACCTCCT | AGGAGGGTCC  |

FIGURE 3H

|       |             |            |             |            |            |
|-------|-------------|------------|-------------|------------|------------|
| 23601 | AGAAAAGCAG  | CTCCCCAAGC | CTGTGCGCCT  | CGTCCTCAGC | AGATCCACCT |
| 23651 | TCTTCACTAT  | AATAAAAGCC | AGTCTGGGAT  | GCTAATAAGG | CCTGTGCTGG |
| 23701 | AGTTTGTACA  | CAAACCTGCA | GAGAGAAAAC  | CAGTGGGGTC | CTGAACCACA |
| 23751 | GCGTGGTCTT  | GGGACAGCCA | CTGCCTTCCT  | CTGGCCCCGG | AGGGAAGCTT |
| 23801 | TGGGGAAGGG  | GCTGGTGGGA | GTTGTTTGCC  | CCACCCTGGC | CTGCTCTGTG |
| 23851 | TGGAAGGCGC  | ACTCCCCAGA | GGGGTGAGTG  | CCAGGCGCTG | TCCGGGTGCC |
| 23901 | TTGGCTTCAC  | GCTGTACCA  | GGCCTGTCCG  | GGACCACCAT | GTTGGTTTCC |
| 23951 | CGTGAGGCCCT | CCCTCTCATA | AGAGGGCCCT  | TCAGAAAGGT | CGGGACCCCT |
| 24001 | CGTAGTGAC   | AAGCTGACAT | CTGCTCCCTG  | CTGGAGGTGG | CTTGACCCA  |
| 24051 | GGGAGAGCCT  | CATAATGAGG | TGGGGGGCCT  | GGGAGAGGCC | TGGAGGTCCC |
| 24101 | AACTGCAGCT  | TTTCTGTCAT | CTCTTCAGGG  | AGGTGGTTGC | GGTTGGGGGA |
| 24151 | GGATTCTCTG  | AGCTCATCCA | GGAATGTAGG  | CCCCTGATGC | TGGAATTGTG |
| 24201 | CTTAGTGTAG  | GGGGAGAGGG | GGCATATATA  | ATTTGACGTC | CAAATGGGGA |
| 24251 | CATTTTTGAG  | AGTGAAGGG  | GAAGCCATTA  | ATAATTATGC | CAGCACGGCC |
| 24301 | GGGTGCGGTG  | GCTCACGCCT | GTAATCCCAG  | CACTTTGGGA | GGCCGAGGCT |
| 24351 | GGTGATCAC   | AGGGTCAGGA | GATCGAGACC  | ATCCTGGCTA | ACACGGTGAA |
| 24401 | ACCCGCTCTC  | TACTAAAAAT | ACAAAAAATC  | AGCTGGGCGT | GGTGGCGGGC |
| 24451 | ACCTGGAGTC  | CCAGCTACTC | AGGAGGCTGA  | GGCAGGAGAA | TGGCGTGAAC |
| 24501 | CCGGGAGGCA  | GAGCTTGACG | TGAGCCAAGG  | TCACGCCACT | GCACTCCAGC |
| 24551 | CTGGGCGACA  | GAGTGAGACT | CCGTCTCAAA  | AAATAATAAT | TATTATGCCA |
| 24601 | GCATGGTGGC  | TCATGCCTAT | AATCCCAGCA  | CTTTGGGAGG | CCAAGGCAGG |
| 24651 | ATTGCTTGAG  | GCCAGGAGTT | CAAGACCAGC  | CTGGGCAACA | TAGCAAGACC |
| 24701 | CCATCTCTAA  | AAAAAAAAAA | AATTAGCCGG  | GCGTGGTGGT | GGGTGCCTGT |
| 24751 | AGTCCAGCA   | ACTCAGGAGG | CTGAGGTGGG  | AGGATTGCTT | GAGTCTGGGA |
| 24801 | GGTGAGGTT   | GCAGTGAGCT | GAGATTGCAC  | CACTGTACTC | CAGCCTGGGT |
| 24851 | GACAGAGCCA  | GACCCTGTCT | CAAAAAAAAA  | AAAGAAAAAA | AAGTAATAAT |
| 24901 | AATTATGCCA  | GGACAGCAGG | TGGACGGACA  | CCTGGTCCTT | CTGACTCAGA |
| 24951 | GCCTGTGGTC  | CAGCACCCCC | TAGTGGTGGA  | ACAAGCCAGA | CACAGGATAA |
| 25001 | GGATACATTT  | AGTGTCTAGT | TTGTACCTGG  | CAAACAGAGT | GACAAGATTG |
| 25051 | GGCTTAATAC  | TTTCCAGCTA | TAAAATTCTA  | GAATTCTGTG | ACCCAAGTTT |
| 25101 | AATTTGGGGT  | AGAGCTTTTT | AAAAAAAAAA  | TAGAGATGGA | GTCTTGCCAT |
| 25151 | GTTGCCCAGG  | CTGGACTTAA | ACTCCTGGCC  | TCAAGCCATT | TGCCCACCTT |
| 25201 | GGCCTCCGAA  | AGTGCCAGGT | GATTACAGGC  | ATGAGCCACC | ACACCCAGCC |
| 25251 | TCCACGTTTA  | ACTTTGAAAG | AAGATTTTAC  | TTCATCATCA | AGTCCCAATA |
| 25301 | TTTATCCTTG  | ATAGACTGCT | TTGGTTTTTT  | GTTTGTITGT | TTTGAGATGG |
| 25351 | AGTTTCACTC  | TTGTTGCCCA | GGCTGGAGTG  | CAGTGGCGCA | AACTCAACTC |
| 25401 | ACTGCAGTCT  | CCGCCTCTCA | CATTCAAGCA  | GTTCTCTTGC | CTCAGCCTCC |
| 25451 | CAAGTACGTG  | GGATTACAGG | TGCATGCCAC  | CACCACACCG | GCTAATTTTT |
| 25501 | GTATTTTTAT  | TAGAGACGGG | ATTTCAACCAT | TTTGCCCAGG | CTGGTCTCAA |
| 25551 | ACTCCTGACC  | TCAGGTAATC | TGCCCCACCTC | AGCCTCCCAA | AGTGCTGGGA |
| 25601 | TTACAGGCGT  | GAGCCACTGT | GCCCCGCCAT  | AGAGTTTTTT | ATACTTTGGG |
| 25651 | ATAATTGTAG  | AAACTCAGTA | GTAGAGTTAA  | GTGGAGTTGG | TCCTTTTTAA |
| 25701 | AGATATCAAA  | ACCCATTTAC | TGGTTATTTT  | AAAAAGAGAC | ATTTTGGGAG |
| 25751 | GAAAACTAGA  | TATAGAAATC | TGTTGAATAT  | GTGACAGAAT | CCCAAGACTG |
| 25801 | ATAGATGGAC  | TCTGCCCTGT | GAACAAGGCA  | AAGAAAAATG | CAAAATGAAA |
| 25851 | GCCTCTCTAC  | CCAGATCTGC | TGGGGGATGA  | CTGAGGTCAA | CACAGAAGGC |
| 25901 | CCTCAGGCCG  | GGCACGGTGG | CTCACGCCTG  | CAATCCCAAC | ACTTTAGGAG |
| 25951 | GCTGAGGTGG  | ATGGATCGCT | TGAGCCCAGG  | AGTTTGAGAC | CAGCCTGGGC |
| 26001 | AACATGGTGA  | AACCCTGTTT | TTATAGAGAT  | AAAAAAATAC | AAAAATTAGC |
| 26051 | TGGGCGTGTT  | GGCATGTGCC | TGTAGTCTCA  | GCTACTCAGG | AGGCTGAGGT |
| 26101 | GGGAGGATCG  | CTTGAGCCTG | GAAGGCAGAG  | GTTGCAATGA | GCTGAGATTG |
| 26151 | CACCACTGCA  | CTGCAGCCTG | CACGACAGAG  | CGAGACGCTG | TCTCAAAACA |
| 26201 | ACAACAAAAC  | CACACACACA | GAGAGAAGGC  | CCTTGATTAG | GCTGATAGTT |
| 26251 | GGAGGATGTA  | GGGAAGTCAG | CTGGGTCAGA  | CTGTGAGCAG | CTCCAGAGGC |
| 26301 | CGTGCTGGGA  | GGTTTAGACT | TCATCTCTGG  | TCAATGGGGG | GCCACGGAGG |
| 26351 | CGTTGCGGGC  | TGAGACTGGG | GGCTGAGAGA  | CCGGCAAGGA | GCAACTGCCG |
| 26401 | TGATGTAGGG  | AGGCCAGAGG | GAGGCCAAGC  | TTGGGGCAGT | GGGTGAAGGG |
| 26451 | GGCTTTGAGA  | GATGTGGGAT | TCAGATTCTT  | GTGTGTGTGA | GGGAGAGTGT |
| 26501 | CTCCCTGAGT  | GCATATTCTG | ACCCTGAGGT  | CCCTCTGTCC | CTGGTGTCCC |

FIGURE 31

|       |             |             |             |             |             |
|-------|-------------|-------------|-------------|-------------|-------------|
| 26551 | CTGAACAGGA  | AAGAAGGGTT  | TGGTGAAGGG  | GGCAAGAGCC  | CAGAGCTCCC  |
| 26601 | CGGCGTCCAG  | GAAGACGAGG  | CTGCATCCTG  | AGCCCCTGCA  | TGCACCCAGG  |
| 26651 | GCCACCCGGC  | AGCACACTCA  | TCCC CGCCT  | CCAGAGGCCC  | ACCCCTCAT   |
| 26701 | GCAACAGCCG  | CCCCCGCAGG  | CAGGGGGCTG  | GGGACTGCAG  | CCCCACTCCC  |
| 26751 | GCCCCTCCCC  | CATCGTGCTG  | CATGACCTCC  | ACGCACGCAC  | GTCCAGGGAC  |
| 26801 | AGACTGGAAT  | GTATGTCATT  | TGGGGTCTTG  | GGGGCAGGGC  | TCCCACGAGG  |
| 26851 | CCATCCTCCT  | CTTCTTGGAC  | CTCCTTGGCC  | TGACCCATTG  | TGTGGGGAAA  |
| 26901 | CCGGGTGCCC  | ATGGAGCCTC  | AGAAATGCCA  | CCCGGCTGGT  | TGGCATGGCC  |
| 26951 | TGGGGCAGGA  | GGCAGAGGCA  | GGAGACCAAG  | ATGGCAGGTG  | GAGGCCAGGC  |
| 27001 | TTACCACAAC  | GGAAGAGACC  | TCCC CGTGGG | GCCGGGCAGG  | CCTGGCTCAG  |
| 27051 | CTGCCACAGG  | CATATGGTGG  | AGAGGGGGGT  | ACCCTGCCCA  | CCTTGGGGTG  |
| 27101 | GTGGCACCAG  | AGCTCTTGTC  | TATTCAGACG  | CTGGTATGGG  | GGCTCGGACC  |
| 27151 | CCTCACTGGG  | GACAGGGCCA  | GTGTTGGAGA  | ATTCTGATTG  | CTTTTTTGTT  |
| 27201 | GTCTTTTACT  | TTTGTTTTTA  | ACCTGGGGGT  | TCGGGGAGAG  | GCCCTGCTTG  |
| 27251 | GGAACATCTC  | ACGAGCTTTC  | CTACATCTTC  | CGTGGTTCCC  | AGCACAGCCC  |
| 27301 | AAGATTATTT  | GGCAGCCAAG  | TGGATGGAAC  | TAACTTTCCT  | GGA CTGTGT  |
| 27351 | TCGCATTCGG  | CGTTATCTGG  | AAAGTGGACT  | GAACGGAATC  | AAGCTCTGAG  |
| 27401 | CAGAGGCCTG  | AAGCGGAAGC  | ACCACATCGT  | CCCTGCCCAT  | CTCACTCTCT  |
| 27451 | CCCTTGATGA  | TGCCCCTAGA  | GCTGAGGCTG  | GAGAAGACAC  | CAGGGCTGAC  |
| 27501 | TTTGACCGAG  | GGCCATGGAC  | GCGACAGGCC  | TGTGGCCCTG  | CGCATGCTGA  |
| 27551 | AATAACTGGA  | ACCCAGCCTC  | TCCTCCTACA  | CCGGCCTACC  | CATCTGGGCC  |
| 27601 | CAAGAGCTGC  | ACTCACACTC  | CTACAACGAA  | GGACAAACTG  | TCCAGGTCGG  |
| 27651 | AGGGATCACG  | AGACACAGAA  | CCTGGAGGGG  | TGTGCACGCT  | GGCAGGTGGC  |
| 27701 | CTCTGCGGCA  | ATTGCCTCAC  | CCTGAGGACA  | TCAGCAGTCA  | GCCTGCTCAG  |
| 27751 | AGCGGGGGTG  | CTGGAGCGCG  | TGCAGACACA  | GCTCTTCCGG  | AGCAGCCTTC  |
| 27801 | ACTTTCTCTC  | TGGGATCAGT  | GTCCGGCTGG  | CCGACGTGGC  | ATTTGCTGAC  |
| 27851 | CGAATGCTCA  | TAGAGGTTGA  | CCCCACAGG   | GTACGCAGG   | ACTCGGACAC  |
| 27901 | TGCCCTGGAA  | ACATGGATGG  | ACAAGGGCTT  | TTGGCCACAG  | GTGTGGGTGT  |
| 27951 | CCTGTTGGAG  | GAGGGCTTGT  | TTGGAGAAGG  | GAGGCTGGCT  | GGGGGAGAAA  |
| 28001 | CCCGGATCCC  | GCTGCATCTC  | CGCGCTGTG   | GGTGCATGTC  | GCGTGCTCAT  |
| 28051 | CTGTTGCACA  | CAGCTCACTC  | GTATGTCCTG  | CACTGGTACA  | TGCATCTGTA  |
| 28101 | ATACAGTTTC  | TACGTCTATT  | TAAGGCTAGG  | AGCCGAATGT  | GCCCCATTGT  |
| 28151 | CAGTGGGTCC  | ACGTTTCTCC  | CCGGCTCCTC  | TGGGCTAAGG  | CAGTGTGGCC  |
| 28201 | CGAAGCTTAA  | AAAGTTACTC  | GGTACTGTTT  | TTAAGAACAC  | TTTTATAGAG  |
| 28251 | TTAGTGGAAG  | GCAAGTTAAG  | AGCCAATCAC  | TGATCCCCAA  | GTGTTTCTTG  |
| 28301 | AGCATCTGGT  | CTGGGGGGAC  | CAC TTTGATC | GGACCCACCC  | TTGGAAAGCT  |
| 28351 | CAGGGGTAGG  | CCCAGGTGGG  | ATGCTCACCC  | TGTCAC TGAG | GGTTTTGGTT  |
| 28401 | GGCATCGTTG  | TTTTTGAATG  | TAGCACAAGC  | GATGAGCAAA  | CTCTATAAGA  |
| 28451 | GTGTTTTAAA  | AATTAAC TTC | CCAGGAAGTG  | AGTTAAAAAC  | AATAAAAAGCC |
| 28501 | CTTTCTTGAG  | TAAAAAAGAA  | AAAAAAAAGG  | TTTGTGCGTA  | CATTTTCTGC  |
| 28551 | ATCTGGATAT  | ACGTTCTTTC  | TCAGCAGCTG  | GAACAGCTGG  | CTTTGTTGAA  |
| 28601 | TTTTCTGGAA  | GCGTCTGAGG  | CACCCTAAGT  | CCCTGAGCAG  | GACAGTGGTG  |
| 28651 | AGAAGTGGTC  | TTGGCGGAGG  | GAGGGAGAGG  | GAAGGGCTGG  | CTCAGGAGGT  |
| 28701 | GACCGGGCTG  | CAGTCCAGGG  | TACAGCTGAG  | GCTCCTGGGC  | GGGTCCGTGG  |
| 28751 | CCACTCCTTG  | GGAAGAAGCTG | CCTGTTTCAC  | AGGGGCTCAG  | GATGCCAAGG  |
| 28801 | TCTGGTCCGG  | GTAGGAGCCA  | TAGCTGCTGC  | TTTTGGGGCA  | GAGGTCCCTG  |
| 28851 | TGGTGT CACA | GGAGTGCCTG  | TGACACCAGC  | CCAGTGACCT  | CCCATCCCCG  |
| 28901 | CTTAGCCTTG  | GACACTGGTA  | CAGACTTTTG  | GGACCCCA    | CCTCTGTTCC  |
| 28951 | CATGGTACAG  | CCCTCCAGGG  | CAGCGACGAA  | AAGAGTCATC  | CTTAAGGTCA  |
| 29001 | CACAGCCCTG  | AGCTTGAATC  | CAAGCTTTGC  | TACTTAAAAA  | TTGTGTGACC  |
| 29051 | TTTGGCAGGT  | CATTGGAGGA  | GCCTCAGTTC  | CCTTATTGAT  | TTAATGGGAA  |
| 29101 | TGTTCCCGTG  | GGGTGTTTTG  | TTTGTTTGTT  | TGAGATTTTT  | TGAGACTTGC  |
| 29151 | TCTGTACCCC  | AGGCTGGAGT  | GCAATGGCAA  | GATCTCGGCT  | CACTGCAACC  |
| 29201 | TCTTCCTCCT  | GGGTTCAAGC  | GATTCTCCTG  | CCTCAGCCTC  | CCAAGTAGCT  |
| 29251 | GGGACTACAG  | GTGCCCGCCA  | CCATGCCCCA  | GCTAATTTTT  | TGTACTTTTA  |
| 29301 | GTAGAGACGG  | GGTTTCACCA  | TGTTGGCCAG  | GCTGGTCTTG  | AACTCCTGAC  |
| 29351 | CTCAGGTGAT  | CTGCCACCT   | CGGCCTCCCA  | AAGTGCTGGG  | ATTACAGGCG  |
| 29401 | TGAGCCACCG  | CGCCACCTC   | CCCATGGGGT  | TTGAATGCAA  | ACAATGCAA   |
| 29451 | CGTTTTCGTC  | TGCTCTCACA  | CTACAACAGT  | GAACACAGAA  | GACTTCTGTG  |

FIGURE 3J



29501 ACCGGCTGGG CGCGGTGGCT CACGCCTGTA ATCCCAGCAC TTTGGGAGGC  
 29551 TGAGGAAGGC GGATCATGAG GTCGGAGATC GAGACCATCC TGGCTGACAC  
 29601 GGTGAAACCC CGTCTCTACT AAAAATACA (SEQ ID NO:3)

FEATURES:

Start: 3000  
 Exon: 3000-3359  
 Intron: 3360-4259  
 Exon: 4260-4307  
 Intron: 4308-4749  
 Exon: 4750-4803  
 Intron: 4804-5206  
 Exon: 5207-5258  
 Intron: 5259-5511  
 Exon: 5512-5645  
 Intron: 5646-6122  
 Exon: 6123-6159  
 Intron: 6160-7038  
 Exon: 7039-7060  
 Intron: 7061-8252  
 Exon: 8253-8341  
 Intron: 8342-12264  
 Exon: 12265-12464  
 Intron: 12465-15226  
 Exon: 15227-15280  
 Intron: 15281-16061  
 Exon: 16062-16135  
 Intron: 16136-18797  
 Exon: 18798-18885  
 Intron: 18886-19070  
 Exon: 19071-19199  
 Intron: 19200-22830  
 Exon: 22831-22934  
 Intron: 22935-26558  
 Exon: 26559-26628  
 Stop: 26629

CHROMOSOME MAP POSITION:

Bac accession number: AC005940  
 Chromosome: 17

ALLELIC VARIANTS (SNPs):

| DNA      |       |       |                | Protein  |       |       |
|----------|-------|-------|----------------|----------|-------|-------|
| Position | Major | Minor | Domain         | Position | Major | Minor |
| 2082     | G     | A     | Beyond ORF(5') |          |       |       |
| 2748     | -     | C A   | Beyond ORF(5') |          |       |       |
| 8868     | G     | A     | Intron         |          |       |       |
| 10187    | G     | A     | Intron         |          |       |       |
| 10216    | T     | C     | Intron         |          |       |       |
| 11100    | G     | A     | Intron         |          |       |       |
| 11650    | G     | A     | Intron         |          |       |       |
| 11656    | A     | C     | Intron         |          |       |       |
| 15095    | C     | T     | Intron         |          |       |       |
| 15116    | A     | G     | Intron         |          |       |       |
| 15428    | A     | G     | Intron         |          |       |       |
| 15827    | T     | C     | Intron         |          |       |       |
| 16135    | G     | A     | Exon           | 375      | G     | E     |
| 16557    | A     | G     | Intron         |          |       |       |

FIGURE 3K



|       |   |       |                |
|-------|---|-------|----------------|
| 17375 | C | T G   | Intron         |
| 17375 | G | C     | Intron         |
| 17375 | A | C G   | Intron         |
| 17375 | A | C G   | Intron         |
| 17511 | C | T G   | Intron         |
| 17928 | A | C     | Intron         |
| 17968 | C | G     | Intron         |
| 18574 | T | C     | Intron         |
| 19654 | C | T     | Intron         |
| 21498 | C | T     | Intron         |
| 22729 | G | A     | Intron         |
| 22757 | C | T     | Intron         |
| 22779 | C | T     | Intron         |
| 24350 | T | G     | Intron         |
| 24558 | T | C A   | Intron         |
| 24872 | - | A     | Intron         |
| 25756 | C | T     | Intron         |
| 25968 | G | A     | Intron         |
| 26537 | G | C     | Intron         |
| 28204 | C | T A G | Beyond ORF(3') |

Context:

DNA  
Position  
2082

ATGCAGAGTCCAGCGCAAGCAGGGGGAAGGGCATCAGGTTGGGCATGGCCAGCGCTCTAC  
AAGCCTGGGACAGAGATGGGGGTCTCAGGCTGAGTGTGAGGGTTTCAGTCCGGGGTCAGGA  
TGTAGCCCAGGGTCATGGCTGAAGGTGAGGGCTGGGGTCACCTCCCTGATGTTTCAGCC  
GCCACACAGTGAGTTTGAGAACATGAGTCTCAGGGGATGTCATGCCCTGTTTCACCCCT  
CATTCCTCATTCCTATCCCTTGCTTTTTTTGAAACCGAGTCTTGCTCCATCACCCA  
[G,A]  
GCTGGAGTGTAGTGGCGTGATCTTGGCTCACTGCAACCTCCACCTCCCAAGTTCACACGA  
TTCTCCTGCCTCAGCCTCCCGAGTAGATGGGATTTTCAGGTGCACGCCACCATGCCTGGCT  
AATTTTTGTATTTTTAATAGAGACAGAGTTTGGCCATGTTAGCCAGGCTAGTCTCGAACT  
TCTGACCTCAGGTGATCCACCTGCCTCGGCCTCCCAAGTGCTGGGATTACAAGTGTGAG  
CCACCATGTGGGGCCATCCCTTGTTTTGACAGACGTCAATGAGGCAGGGCTGGCTGGA

2748

CAAGGAGGGCTCCCTGGAGGAGGCGGTGGGTCTGAAGCATCAGCAAGGCTTCTGAGTTA  
CTAGTGTCTAGCTCAGCTTCCAGGAGGCAGTGTGCGAGTGCTCTGCTGTCAAGGGTTGGG  
ACTCATGACTACAGGGCTGCATGCTGTGCTGGGGCTGAGCTGACCCTGGGCTCTGCCCC  
TTCCAGTGCTGCTGGGCCTCCAGGCTTCTGCCCTGTCTGCTCCTGATTCCAGAATATCAGA  
TTCTCTGCTTCCCTGTGAAGCCAGCAGGCAGAAAGTGAAGTGCCTCTGTTACCGGCAGGG  
[- ,C,A]  
TACTGAGGCCTAGAGGGCTGGCATGCGGCAGAACCGATGTGAATTCATTGAGTTCATAGG  
GACAGACTTGAGTTTGGGTGTTGGCAATCCCGGTAGAGGGAACAGCCAGGGCAAAGGCAT  
GGAGGTGGGACCCACAGCGCTGTGGCTACCTTACCTGGTAGCCAGCCTGACACCCAGGAG  
TGAAGCCTTCTCTGCCTTCTTTTCTCAGGTTCCCAACAAGGCTACGCAGAAGAACCCCT  
TGACTGAAGCAATGGAGGGGGTCCAGCTGTCTGCTGCCAGGATCCTCGGGCAGAGCTGG

8868

TCCTGTGGGTCTGAGTCCTGGCTGTTCTGTATCTTTCTTCTGCTGAGTTCCTAGCCTAGC  
TTAGCGTTGCCACGGGGCTTCAAGAGATGTGGGAAGGAAGGGATTTATGTCCAGCTGCTG  
GGGAGAGTCTGTCTGGCATGGGGCCGGGCATGGTGGCAGGGTGGATTTACCTGTGAGG  
GGCCCTAGTCTGATAAGAGCTCAGGAGGGTGTGTGAGCTTGGCCTCTGTCTCATTTTCAT  
TCATTAGCTACATTCATTGCCTGGGGGCATAGGGGTGAAAGACCCAGACCCGAGTTCAC  
[G,A]  
GCCTAGTGGGAGGGACAGGAATCTAGGCAGGCAGATAATACAGCGTGGTGCCTGCCAAGG  
CTGGGAGCCTAGAGGCTGTAGGAGTGCCGGGGGGCTGGGGAAGTCTCCCTGAAGAGGCT  
ACTTATGATTGGGTCTGAGGGATGAGTAGACTTCCCTGCTCAGGTTTTGAGGGATGGG  
CGTGGAAGACGATGTGCTGGCATAGGCGTGTACTCTGAGTCTGGGAGAAGTGGAGTCT

FIGURE 3L

GGCTGAAGCCTCCAGTGGGCAGAGGAGGGCCGTGGTTAGTGAAAGATGATGCTGGAAACA

10187 TTCCCTGGGGGAACAGGGTGGGGATCTGGCTGAGGTTTGACCTGCAGTGACAGAAACAGG  
 ACTGTCTTTATCCTGCTCGAGCCTCTCCTTTGCCTTCAGATTAAGACTCTCTTTGCACAT  
 ATGGGGAAACTGAGGCACACAGAGGGGAGGGCTTTGCAGAAAATCCCTACCAAGGGCCTA  
 GAGGCATGGGATGGGAAGGGGACATTTTACCCCGTACGGTCAGTGGCAGGCAC  
 [G,A]  
 GTCCTGTACCAGCTTGGCTCCACCTCCTTTCTGTTGTAGTCCCTTCTTTCCCTGAAGTC  
 CTGTTGTCTGCTATCCCCTAGCCTCCACAAAGAAACGAGTTTATCTTACCTGGTTCTTG  
 GTAAAGCCTCATCAGGACCCAGCTAATCACAGTGAAGGGCTTCCCTGGGGCAGAACGGTT  
 AGCGCCAGGGGCTGGACAGTGGATGAACAGAGGCACGAGGGCGCTGAAGACCTGCCTTG  
 TGATTCTGGCCCCAAGAAGAGAGAGTTGAGGCTGCCATGAGAGGGCTCGGTGGTCAGGGC

10216 TTCCCTGGGGGAACAGGGTGGGGATCTGGCTGAGGTTTGACCTGCAGTGACAGAAACAGG  
 ACTGTCTTTATCCTGCTCGAGCCTCTCCTTTGCCTTCAGATTAAGACTCTCTTTGCACAT  
 ATGGGGAAACTGAGGCACACAGAGGGGAGGGCTTTGCAGAAAATCCCTACCAAGGGCCTA  
 GAGGCATGGGATGGGAAGGGGACATTTTACCCCGTACGGTCAGTGGCAGGCACAGTCCT  
 GTACCAGCTTGGCTCCACCTCCT  
 [T,c]  
 TCTGTTGTAGTCCCTTCTTTCCCTGAAGTCCTGTTGTCTGCTATCCCCTAGCCTCCACA  
 AAGAAACGAGTTTATCTTACCTGGTTCTTGGGTAAAGCCTCATCAGGACCCAGCTAATCA  
 CAGTGAAGGGCTTCCCTGGGGCAGAACGGTTAGCGCCAGGGGCTGGACAGTGGATGAAC  
 AGAGGCACGAGGGCGCTGAAGACCTGCCTTGTGATTCTGGCCCCAAGAAGAGAGAGTTGA  
 GGCTGCCATGAGAGGGCTCGGTGGTCAGGGCGGCCAGGCCTGGTTCTCAGTTGATGGGG

11100 ATCGATGCCTAGCTGCCTAGTTATCTACCAAGATGTCTATTGGAAATCTATGTGGGTAAA  
 GAGCTGGAGTTCAAGGGAGAGGCTAGGGTTTGAGATAAGAACATGAGACCACTTTCCATG  
 GTCAAATGTCCACCCCCCTGAGCTTCTGTGCCCTGAAGGGTGTGTGAGATTCTTTGTGTG  
 TGCCTGGCACATAGTAGGCAATCAAGAAAGTGCCACTGGTTTTATGGTTATTGTTATACG  
 GCACCCGCCCTTCTCTGCCCGCAGCCTCCCTCTCCTCTTCTCCCTTCTCTTTCTTCTCTC  
 [G,A]  
 CCTTCTCTCCTCCCTCCTCTCCAGCATCCTGGGGTCCGTTGGTCCAGATGAAGGTACTTG  
 CCAAGGAGGGAGCCACAGTGCATGGTCCGCGGATGGGGTCAGTGGGGTCATTGTCTCT  
 CTTGGCTGGGACCTTACCAGTCATGTGAGCTTGAGCCACCTGTCACTTCGTGGTGGTGCT  
 GGGCCAGAAAGCAGGGCAGACCTCCAGCCTATTAGGTCACTTCTGATTGGGATTCGTC  
 CTAATATATGTGGCTGACCTTACACCCAGCTGTGTATCCTGCTTGTCCCAAGGCCTGG

11650 GTGGCTGACCTTACACCCAGCTGTGTATCCTGCTTGTCCCAAGGCCTGGGGTGCCATC  
 CATCTCTGTGAAACCCCATCAGCCAGATCCCGAGGGCTGAGATGGTACCTCTGTAGGAT  
 AGCAGAGTCCCTACAATCTTACTCTAGTCCCAGCAGCAGGGACATCTTTGCCTAGCCTG  
 GGTGGGGGATGGAAGTGGAGAAAGGTTTTGATTGGCTTTGGGCCTGCAGACGGCACTCAC  
 AGGGAAGGGGCAGAGCTAGCCTAGGAAGAACTCTGCTCCAGCTGGGGGCGGTGGCTCAC  
 [G,A]  
 CCTGTAATCCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCACCTGAGGTGAGGAGTTCA  
 AGACCAGCCTGACCAACATGGCGAAACCCTGTCTCTACTAAAAATACAAAAAGTAGCCGG  
 GCGTGGTGGCAGACACCTGTAATCCCAACTACTCGGGAGGCTGAGGCAGGAGAATCTCTT  
 GAACCTGGGAGGTGGAGGCTGCAGTGAGCCGAGATCACGCCATTGCACTCCAGCCTGGGG  
 GACAGAGTGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAACCAAAAAAAAAAACAGCAACA

11656 GACCTTACACCCAGCTGTGTATCCTGCTTGTCCCAAGGCCTGGGGTGCCATCCATCTC  
 TCTGAAACCCCATCAGCCAGATCCCGAGGGCTGAGATGGTACCTCTGTAGGATAGCAGA  
 GTCCCTACAATCTTACTCTAGTCCCAGCAGCAGGGACATCTTTGCCTAGCCTGGGTGGG  
 GGATGGAAGTGGAGAAAGGTTTTGATTGGCTTTGGGCCTGCAGACGGCACTCACAGGGAA  
 GGGGCAGAGCTAGCCTAGGAAGAACTCTGCTCCAGCTGGGGGCGGTGGCTCACGCCTGT  
 [A,c]  
 ATCCAGCACTTTGGGAGGCCGAGGTGGGTGGATCACCTGAGGTGAGGAGTTCAAGACCA  
 GCCTGACCAACATGGCGAAACCCTGTCTCTACTAAAAATACAAAAAGTAGCCGGGCGTGG  
 TGGCAGACACCTGTAATCCCAACTACTCGGGAGGCTGAGGCAGGAGAATCTCTTGAACCT  
 GGGAGGTGGAGGCTGCAGTGAGCCGAGATCACGCCATTGCACTCCAGCCTGGGGGACAGA

FIGURE 3M

GTGAGACTCTGTCTCAAAAAAAAAAAAAAAAAACCAAAAAAAAAACAGCAACAACTCTC

15095 GGCTGTTTGCAATATCAGGAAGAAGGTGGATTATGAGGAGAAGGGATGACTCCTTGAAGC  
CCGAGCTGGTTTAGTGAGCAGAAGTTCATATATACCATCATTCTGCGGTGCGTCTGTG  
GCACGGGAGCGGCCGTGTGACCCTCTGGATGAAGGAGGTTTGTACCTGTTGAGTTGGA  
AACGTACCTGGTTAGAGTCTTTCCCAAGGAAACCCAGAACCCCTGGAGGGTGGAGGCCTT  
GTTCTGGCCGCCCTGTGCTCCTCAGCACTCAGCACGGGGCCCAGCATCGGGCAAGTACCG  
[C, T]  
GGAGTGTTTGTCGAGTGAGTACATGACAGAGGAAAGAGGTTCCCTGCAGGCCTCTCCTGC  
AGCCCGCTGGAGCTGGGTGGCAGAGGTGGCTGTGCCTGTTGGGGACTGATGTGAGCATG  
TTTCTTTCCAGGCCTTGGATGTATGGGCCACTGGCGTCACGTTGTACTGCTTTGTCTATG  
GGAAGGTGAGTGCCAGGGATGCCAGCAGAGCTGGGGCGGGTCCAGTGAGGCGGGCACGGG  
CGACGGATGCAGGCTCTTCTTTTTGTCTTAAGTGGCTTTTGAAAGAGCCACCTGGCT

15116 GAAGGTGGATTATGAGGAGAAGGGATGACTCCTTGAAGCCCGAGCTGGTTTAGTGAGCAG  
AAGTTCATATATACCATCATTCTGCGGTGCGTCTGTGGCACGGGAGCGGCCCGTGTGA  
CCCTCTGGATGAAGGAGGTTTGTACCTGTTGAGTTGGAACGTACCTGGTTAGAGTCTT  
TCCCAAGGAAACCCAGAACCCCTGGAGGGTGGAGGCCTTGTCTGCGCCGCCCTGTGTCC  
TCAGCACTCAGCACGGGGCCCAGCATCGGGCAAGTACCGCGGAGTGTTTGTGAGTGAGT  
[A, G]  
CATGACAGAGGAAAGAGGTTCCCTGCAGGCCTCTCCTGCAGCCCGCTGGAGCTGGGTGGG  
CAGAGGTGGCTGTGCCTGTTGGGGACTGATGTGAGCATGTTTCTTTCCAGGCCTTGGATG  
TATGGGCCACTGGCGTCACGTTGTACTGCTTTGTCTATGGGAAGGTGAGTGCCAGGGATG  
CCAGCAGAGCTGGGGCGGGTCCAGTGAGGCGGGCACGGGCGACGGATGCAGGCTCTTCT  
TTTTGTCTTAAGTGGCTTTTGAAAGAGCCACCTGGCTCAGAGAAGGCTGAGAGAGAAG

15428 AAAGAGGTTCCCTGCAGGCCTCTCCTGCAGCCCGCTGGAGCTGGGTGGGCAGAGGTGGCT  
GTGCCTGTTGGGGACTGATGTGAGCATGTTTCTTTCCAGGCCTTGGATGTATGGGCCACT  
GGCGTCACGTTGTACTGCTTTGTCTATGGGAAGGTGAGTGCCAGGGATGCCAGCAGAGCT  
GGGGCGGGTCCAGTGAGGCGGGCACGGGCGACGGATGCAGGCTCTTCTTTTTGTCTTA  
AGTGGCTTTTGAAAGAGCCACCTGGCTCAGAGAAGGCTGAGAGAGAAGAGGCTTTTTCT  
[A, G]  
TCTTTCTCTGGTCCCCTGCGGAGCGATTCTCGGAAGGAGTCGAGGACAGCAGACACCT  
AAGGGGAGGTGCCGACGATGGTGTGTCACCGCCCCAGCCAGAGTGCTCCCCGTCCCTCT  
GTCCCTTGACGCCATTCACTTATTGAGCCATGTGTTCACTCCCTTGCTCATTTATTCGAC  
AAATTGTCTTACCCCTACCCTGGCTGAGGCTGGACCCTGGGGACACCCAACGCTGACG  
TATCGGTGATCCCTGCCCCAGGTGTGCCTGCTCTGGTGACCACACTAAGGGGCAGGGGG

15827 CCAGAGTGCTCCCCGTCCCTCTGTCCCTTGACGCCATTCACTTATTGAGCCATGTGTTCA  
CTCCCTTGCTCATTTATTGACAAATTGTCTTACCCCTACCCTGGCTGAGGCTGGACC  
CTGGGGACACCCAACGCTGACGTATCGGTGATCCCTGCCCCGAGGTGTGCCTGCTCTGGT  
GACCACACTAAGGGGCAGGGGGGAATTTAGTGAACATGTTCCAAGCCCCAGGCCCTGG  
GAGTGAGGCCTGGCCACAGGTGGCGGTAAATGGTGGTGGGTGCACCCAGCCTGGCTGGC  
[T, C]  
TGGCCGCGGGTGGCAGTAACGGCGGTGGATGCACCCAGCCTCATTGTTCCCTCAGCAACT  
CATTCATTAGTCAACATTTGTTGAACATTTACAGTGTGAGTTGAGGTCTTCTCATGTA  
ATGGGAGCCCAGACCTGCCCCCTACCCCTGCCCCACCAAGGGAGGGGGGTTGATCCCC  
GGCAGAGTGCAGGCCCTGGACCCACATCCTTTGTCTGCCTCTCCACCCACAGTGCCCG  
TTCATCGACGATTTATCCTGGCCCTCCACAGGAAGATCAAGAATGAGCCCGTGGTGTTC

16135 GGGTGGCAGTAACGGCGGTGGATGCACCCAGCCTCATTGTTCCCTCAGCAACTCATTAT  
TCAGTCAACATTTGTTGAACATTTACAGTGTGAGTTGAGGTCTTCTCATGTAATGGGAG  
CCCAGACCTGCCCCCTACCCCTGCCCCACCAAGGGAGGGGGGTTGATCCCCTGGCACAG  
GTCGAGGCCCTGGACCCACATCCTTTGTCTGCCTCTCCACCCACAGTGCCCGTTATCG  
ACGATTTATCCTGGCCCTCCACAGGAAGATCAAGAATGAGCCCGTGGTGTTCCTGAGG  
[G, A]  
GTGAGTTGTCCACCCAGGGGAACAAGGGGGCTACCACCCGCTCCTGGTGTCTGAGTTTFA  
GCAGAGCTTTTGCCCTCTGAGGACCCACCCAGCCTGCAGATATGAAGGTGGCGGTGCT  
GTTCCCTGGGAGGGACCCCTGAATAGATGGACGGGAGGACTCTGGAGCCAAGGTCTCC

FIGURE 3N

GCAACGTCACTGTGTGGATGGGAACCTGAGATCCAGGGTTGGCCAGGGATGACCACAGG  
CATCATTACACCACTCCTTCACCGCAGGCCTGCCTGGGGTCAGTGGCGCCAGCCCCACC

16557 TTCCCTGGGAGGGACCCCTGAATAGATGGACGGGAGGGACTCTGGAGCCAAGGGTCTCCG  
CAACGTCACTGTGTGGATGGGAACCTGAGATCCAGGGTTGGCCAGGGATGACCACAGGC  
ATCATTACACCACTCCTTCACCGCAGGCCTGCCTGGGGTCAGTGGCGCCAGCCCCACCC  
AGCCCCCTGGACTCAAGGGGAACCTTCTCCTTCCCCACTCAGGGTCAGGGAACCTTCAAGAT  
GCCAGTGCCTGCTCCCCATTTACAGATGGAAGAGGATGCTCTGGAGGAGAGCGGTCA  
[A, G]  
GGGGCTGGGACTCAAGCCACTCTTCTCCCCACTCTTCCATTGTGACCGAGGTCTCTGA  
GCGTAGCAGGGATGTGCGGGAGGCCTCTTGCTCATGCATGGTTCGCTCATGACGGCCAC  
CGTGGCAGCCACAGCCTGAGCTCCCAGGCTCCTCTTTTCAGCAGTGGATTTAGGAGTGA  
AATGGAGGCCGGGTGCGGTGGCTCACGCCTGTAAT

17375 GAGGCCAAGGAGGGCGGATCACGAGGTGAGGATCGAGACCATCCTGGCTAACACGGTG  
AAACCCCATCTCTACTAAAAATAGAAAAAAATTAGCCGGGCGTGGTGGCGGGCGCCTGT  
AGTCCCAGCTACTCGAGAGGCTGAGGCAGGAGAACGGCGTGAACCCGGGAGGCAGAGGTT  
TCAGTGAGCTGAGACCGTGCTACTGCACTCCAGCCTGGGCGACAGAGCGAGACTCTGTCT  
CAAA  
[C, T, G]  
AAAATGAAACGGGACTTGTACTCAGCGACTCCTGCTCTCTTCTGCTTATTTCTGTGTGG  
TCCCCAAGCCCTGCTGAGCCCTCCTCTTCCCTGTCTCTGGGCTTGTGCCACTTATACC  
CCTTGCTCATTAGGCCTCAGGCCCCCTCCCAGACTTATCTAGCCACCTTCCCCCTGGT  
CTCGCTGCTGCTGGCCTCCCTCCAGTCCAGCCAACACATTAGGCGGGGACAGCCCTGAT  
AAAGCACAACAATCTGCCTGCATCTCTTGCTGAAGTTTGTCTGAAGCTTCTCAAAGCC

17375 GAGGCCAAGGAGGGCGGATCACGAGGTGAGGATCGAGACCATCCTGGCTAACACGGTG  
AAACCCCATCTCTACTAAAAATAGAAAAAAATTAGCCGGGCGTGGTGGCGGGCGCCTGT  
AGTCCCAGCTACTCGAGAGGCTGAGGCAGGAGAACGGCGTGAACCCGGGAGGCAGAGGTT  
TCAGTGAGCTGAGACCGTGCTACTGCACTCCAGCCTGGGCGACAGAGCGAGACTCTGTCT  
CAAA  
[G, C]  
AAAATGAAACGGGACTTGTACTCAGCGACTCCTGCTCTCTTCTGCTTATTTCTGTGTGG  
TCCCCAAGCCCTGCTGAGCCCTCCTCTTCCCTGTCTCTGGGCTTGTGCCACTTATACC  
CCTTGCTCATTAGGCCTCAGGCCCCCTCCCAGACTTATCTAGCCACCTTCCCCCTGGT  
CTCGCTGCTGCTGGCCTCCCTCCAGTCCAGCCAACACATTAGGCGGGGACAGCCCTGAT  
AAAGCACAACAATCTGCCTGCATCTCTTGCTGAAGTTTGTCTGAAGCTTCTCAAAGCC

17375 GAGGCCAAGGAGGGCGGATCACGAGGTGAGGATCGAGACCATCCTGGCTAACACGGTG  
AAACCCCATCTCTACTAAAAATAGAAAAAAATTAGCCGGGCGTGGTGGCGGGCGCCTGT  
AGTCCCAGCTACTCGAGAGGCTGAGGCAGGAGAACGGCGTGAACCCGGGAGGCAGAGGTT  
TCAGTGAGCTGAGACCGTGCTACTGCACTCCAGCCTGGGCGACAGAGCGAGACTCTGTCT  
CAAA  
[A, C, G]  
AAAATGAAACGGGACTTGTACTCAGCGACTCCTGCTCTCTTCTGCTTATTTCTGTGTGG  
TCCCCAAGCCCTGCTGAGCCCTCCTCTTCCCTGTCTCTGGGCTTGTGCCACTTATACC  
CCTTGCTCATTAGGCCTCAGGCCCCCTCCCAGACTTATCTAGCCACCTTCCCCCTGGT  
CTCGCTGCTGCTGGCCTCCCTCCAGTCCAGCCAACACATTAGGCGGGGACAGCCCTGAT  
AAAGCACAACAATCTGCCTGCATCTCTTGCTGAAGTTTGTCTGAAGCTTCTCAAAGCC

17375 GAGGCCAAGGAGGGCGGATCACGAGGTGAGGATCGAGACCATCCTGGCTAACACGGTG  
AAACCCCATCTCTACTAAAAATAGAAAAAAATTAGCCGGGCGTGGTGGCGGGCGCCTGT  
AGTCCCAGCTACTCGAGAGGCTGAGGCAGGAGAACGGCGTGAACCCGGGAGGCAGAGGTT  
TCAGTGAGCTGAGACCGTGCTACTGCACTCCAGCCTGGGCGACAGAGCGAGACTCTGTCT  
CAAA  
[A, C, G]  
AAAATGAAACGGGACTTGTACTCAGCGACTCCTGCTCTCTTCTGCTTATTTCTGTGTGG  
TCCCCAAGCCCTGCTGAGCCCTCCTCTTCCCTGTCTCTGGGCTTGTGCCACTTATACC  
CCTTGCTCATTAGGCCTCAGGCCCCCTCCCAGACTTATCTAGCCACCTTCCCCCTGGT  
CTCGCTGCTGCTGGCCTCCCTCCAGTCCAGCCAACACATTAGGCGGGGACAGCCCTGAT  
AAAGCACAACAATCTGCCTGCATCTCTTGCTGAAGTTTGTCTGAAGCTTCTCAAAGCC

FIGURE 30

CTCGTGCTGCTGGCCTCCCTCCAGTCCAGCCAACACATTTCAGGCGGGGACAGCCCTGAT  
AAAGCACAACAATCTGCCTGCATCTCTTGCTGAAGTTTGTCTGAAGCTTCTCAAAGCC

17511 GAGGCTGAGGCAGGAGAACGGCGTGAACCCGGGAGGCAGAGGTTTCAGTGAGCTGAGACC  
GTGCTACTGCACTCCAGCCTGGGCGACAGAGCGAGACTCTGTCTCAAAAAAAAAAAAAA  
AAAAAAAAAACAAAAAAAAAAAAACAAAAACAAAAAGAAAAATGAAACGGGAC  
TTGTACTCAGCGACTCCTGCTCTCTTGCTTATTTCTGTGTGGTCCCCAAGCCCTGCT  
GAGCCCTCCTCTCCCTGTCTCTGGGCCTTGTTGCCACTTATACCCCTTGCCCTATTAG  
[C,T,G]  
CCTCAGGCCCCCTCCCCAGACTTATCTAGCCACCTTCCCCCTGGTCTCGCTGCTGCTGGCC  
TCCCTCCAGTCCAGCCAACACATTTCAGGCGGGGACAGCCCTGATAAAGCACAACAATCT  
GCCTGCATCTCTTGCTGAAGTTTGTCTGAAGCTTCTCAAAGCCACACCCTGGCGTAGC  
ATTACACGTCTCCGGTTCTGCCACCCGCTCGTCTGGGGCCGCTCACTCCCTTTCCCG  
AGCACCAGCCAGCTGGCTTCTGTCCATTTCTCCTCATCCTGTGGTTGCCTTCCCTCCCT

17928 ATCTGCCTGCATCTCTTGCTGAAGTTTGTCTGAAGCTTCTCAAAGCCACACCCTGGCGC  
TAGCATTACACGTCTCCGGTTCTGCCACCCGCTCGTCTGGGGCCGCTCACTCCCTTT  
CCCGAGCACCAGCCAGCTGGCTTCTGTCCATTTCTCCTCATCCTGTGGTTGCCTTCCCT  
CCCTGCCTCCACAGTTGTACCCCTGGTGCCTCTCTTCTGTCTATACCCCTGCTGAGGGG  
TGTCTTTCCCTCAGCCAGGAATTTAAAAGGGATGAAGCATCTAAGACAACAGGGGGA  
[A,C]  
CCGAAGTCAACAGTCTGAGAGTGGCTTTCTGCTCCCTACTCTTGGAAGGATGGGCTCCC  
CAAGACCACTGGTGGCAAAGAAACCTGGGGTTTGGCCGGGCGTGGTGGCTACGCCTGTA  
ATCCCAGCACTTTGGGAGGCCAAGGCAGGCGGATCATGAGATCAGGAGATCGAGATCATC  
CTGGCTAACACGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCCGGGCACGGT  
GGCGGGCACCTGTAGTCCAGCTACTCGGGAGGCTGAGGCAGGAGAATGGCATGAACCTG

17968 TCAAAGCCACACCCTGGCGTAGCATTACACGTCTCCGGTTCTGCCACCCGCTCGTCT  
GGGGCCGCTCACTCCCTTTCCCGAGCACCAGCCAGCTGGCTTCTGTCCATTTCTCCTC  
ATCCTGTGGTTGCCTTCCCTCCCTGCCTCCACAGTTGTACCCCTGGTGCCTCTCTTCTG  
CTATACCCCTGCTGAGGGGTGTCTTTCCCTCAGCCAGGAATTTAAAAGGGATGAAG  
CATCTAAGACAACAGGGGAACCGAAGTCAACAGTCTGAGAGTGGCTTCTGTCCCTA  
[C,G]  
TCTTGGAAGGATGGGCTCCCAAGACCACTGGTGGCAAAGAAACCTGGGGTTTGGCCGGG  
CGTGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCAGGCGGATCATGAG  
ATCAGGAGATCGAGATCATCCTGGCTAACACGGTGAAACCCCGTCTCTACTAAAAATACA  
AAAAATTAGCCGGGCACGGTGGCGGGCACCTGTAGTCCAGCTACTCGGGAGGCTGAGGC  
AGGAGAATGGCATGAACCTGGGAGGCGGAGCTTGCACTGAGCCGAGATTGCGCCACTGCA

18574 GCCTGGGCCACAGAGCGAGACTCCATCTCAAAAAAAAAAAAAAGAACCTGGGGTTTGGGCA  
GAGAGAGTTGGAGCTGATGTGGCGCTGAGGGGGCTGCTCCCTCCCATCTGAGTCTCCCAT  
CTCTGCCTGCACTCTTCTGGCTGGCACTGTGCCAGCTGCTAACCTCCCTGGGCCTCAGT  
TTCCTCCTCTGTCAAATGAGAGAGGATCTTCTCTGGGTGTAGAAAAGGACGAGGTGGTGA  
GTGGGTCTGAAGGCCTCTGGTGTCCATAAAGCGACTCTCCTCACCATCTTTGCCACCCA  
[T,C]  
TGGGGTGTCCAGCACCCATGGAACCTGTCTGTGCCTCTGTCTGGAGGGAGACTTGACC  
TCCTGCTCAGGAAAGGCTCTCAAGCCCTTGTGTGAAATTCCTGCCTGCTGTCCGGAAC  
TCAGTCTTCCCATCCGAGGGACGAAGTTTCGGGAAGAGAGGTGGACAGGAAGGGTCTCT  
CATCAGCGGTCCCACCTCCTCTCCTTCTTCCCTCTCCAGGCCAGAAATCAGCGAGG  
AGCTCAAGGACCTGATCCTGAAGATGTTAGACAAGAATCCCAGACGAGAATTGGGGTGC

19654 AGAGGGAGGTCTTGTGGTGGGAGACCAGGAGGCTTGGTGAGGAGAGTGAATTTAAA  
GAAATAGCGGGCGTGGGGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGA  
GGCCAAGGCGGGCAGATCACGAGGTGAGGAGATCGAGACCATCCTTGAAACCCGACTCT  
ACTAAAAATATAGAAAAATTAGCTGGGCGTGGTGGCGGGCGGTGTAGTCCCAGCTACTCG  
GGAGGCTGAGGCAGGAGAATGGTGTGAACCCGGGAGGTGGAGTTTGCCGTGAGCCGAGAT  
[C,T]  
GCGCCACTGCACTCCAGCCTGGGCCACAGAGCGAGACTGCGTCTCAAAAAAAAAAAAAAGA  
AGAAAAGAAAAGAAAATACCGGGCGCGGTGGCTCACGCCTGGAATCCCAGCACTTTG

FIGURE 3P

GGAGGCCGAGGCGGGTGGATCACGAGGTCAGGAGATCGAGACCATCCTGGCTAATACGGC  
GAAACCCACCTCTACTAAAAATACAAAAAATTAGCCGGGCGCAGTGGTGGGCACCTGT  
AGTCCCAGCTACTGGGGAGGCCGAGGCAGGAGAATCGCTTGAACCTGGGAGGTGGAGGTT

21498 GGTGCTGGAACCAGGCAACTCCCACGGGGTCCCCATGACCACTTGCTGATCTTAGCCA  
CCATCTCCTCTCTCAGACCACTGGAACAACCTCCCACGCTGTCCCTTGCTTCTACTCT  
CACTCCCTGTCCCCCTGGTCAATGCTCAACTCAGCACCCAGCATGGTCCCAGTGGCATGA  
GTGTGTACCTCCCAGCTCAGAGCCTGCTTCTCACTCGGGCTGCTGTGTCCCTCAGAATC  
AGACCTCCAGCCTGTGCCCCACACCCGCCCTGTTTTCTGCGGGGCTCGTGACCGTCC  
[C,T]  
GCCATCATGCACTCGTCTCTGGCCACGTGCCATGGAAGGGGCTGCCCCAGAGCCTTCAGA  
CTTCGCTTCCCTCTGCCCCGGGAGTCCCACCCCGATGGCCACGGGACTCGTCCCTCAC  
TTCTTTCGGCTTTTACGCCAGGGTCCCCTCCTAGAGAGAAGCGAGCCTTCCCTGACCCT  
GTAGCTTCAGCCTCCCCTGCTTCACACCTCATCGCCATTCCCTTGTTTTATTTTTCTT  
TCCACTTACTGACATACATAATTTACTGATTTTCTTCTTACTTATCGCCTGTCTCCCC

22729 GCCCTGCAGTAGCATCTTGGCATCTTCTCGGCGGCCGGAAGGCGGGAAGGATGGCACAGC  
ATCCTTCCATGGCATTGCTGCCGTAGCGAGAAGGTATCTTCTAATGGACTCCCACTTCCA  
GCCCTGGCCCTCCCACTCTTTCAGCCTGGCCTTGGCGACCTTCATGGGCTGGTCCCGG  
CCCCCTCCTCATGTACCACTGGCATCCGGCTCCTCACCATTCCAGGAATATGCCCCAGC  
TGCCAGCGCCCCGTGTTCTTGCTCTGCCATTCATGCTGTGCTGATTGAGATGGGACCC  
[G,A]  
CACTGCGGCCCCCTTGGCAGCTGCTCTCGGGGAATCGGAGCAGAGGCTGCGTGTCTGGGA  
GCCTGGGACCTGTGCTCCTCACGCTGCCTTGCTCCTCAGATCCTGGTGAAGTCCATGC  
TGAGGAAGCGTTTCTTTGGGAACCCGTTTGAAGCCCAAGCACGGAGGGAAGAGCGATCCA  
TGTCTGCTCCAGGAACTACTGGTGAAGTACTGGTGGGCCAGGGACTGCCGGGCACTC  
CCTGGAGTTGGGTGGGAGGTCTGAGGCCCATCCTCCCACTCTCACTGTCTGTTGGGCCAA

22757 CGGCGGCCGGAAGGCGGGAAGGATGGCACAGCATCCCTCCATGGCATTGCTGCCGTAGCG  
AGAAGGTATCTTCTAATGGACTCCCACTTCCAGCCCTGGCCCTCCCACTCTTTCAGCCT  
GGCCTTGGGACCCCTTCATGGGCTGGTCCCAGCCCCCTCCTCATGTACCACTGGCATCCG  
GCTCCTCACCATTCCAGGAATATGCCCCAGCTGCCAGCGCCCCGTGTTCTTGCTCTGC  
CATTTTATGCTGTGCTGATTGAGATGGGACCCGCACTGCGGCCCTTGGCAGCTGCTCT  
[C,T]  
GGGGAATCGGAGCAGAGGCTGCGTGTCTGGGAGCCTGGGACCTGTGCTCCTCACGCTGCC  
TTGTCCTCCTCAGATCCTGGTGAAGTCCATGCTGAGGAAGCGTTTCTTTGGGAACCCGTT  
TGAGCCCCAAGCACGGAGGGAAGAGCGATCCATGTCTGCTCCAGGAAACCTACTGGTGT  
AGTACTGGTGGGCCAGGACTGCCGGGCACTCCCTGGAGTTGGGTGGGAGGTCTGAGGC  
CCATCCTCCCACTCTCACTGTCTGTTGGGCCAAGGCCAGAGCCTGGGACTTGGCCAGGTC

22779 ATGGCACAGCATCCCTCCATGGCATTGCTGCCGTAGCGAGAAGGTATCTTCTAATGGACT  
CCCACTTCCAGCCCTGGCCCTCCCACTCTTTCAGCCTGGCCTTGGGACCCCTTCATGGG  
CTGGTCCCGGCCCTCCTCATGTACCACTGGCATCCGGCTCCTCACCATTCCAGGAATA  
TGCCCCAGCTGCCAGCGCCCCGTGTTCTTGCTCTGCCATTTATGCTGTGCTGATTGA  
GATGGGACCCGCACTGCGGCCCTTGGCAGCTGCTCTCGGGGAATCGGAGCAGAGGCTG  
[C,T]  
GTGTCTGGGAGCCTGGGACCTGTGCTCCTCACGCTGCCTTGCTCCTCAGATCCTGGTG  
AAGTCCATGCTGAGGAAGCGTTTCTTTGGGAACCCGTTTGAAGCCCAAGCACGGAGGGAA  
GAGCGATCCATGTCTGCTCCAGGAAACCTACTGGTGTAAAGTACTGGTGGGCCAGGACTG  
CCGGGCACTCCCTGGAGTTGGGTGGGAGGTCTGAGGCCCATCCTCCCACTCTCACTGTC  
GTTGGGCCAAGGCCAGAGCCTGGGACTTGGCCAGGTCTCGGTGTTGGCCCCATTTGCAT

24350 AGGGAGAGCCTCATAATGAGGTGGGGGCTGGGAGAGGCCTGGAGGTCCCAACTGCAGC  
TTTTCTGTATCTCTTACGGGAGGTGGTTGCGGTTGGGGGAGGATTCTCTGAGCTCATCC  
AGGAATGTAGGCCCTGATGCTGGAATTGTGCTTAGTGTAGGGGGAGAGGGGCATATAT  
AATTTGACGTCCAATGGGGACATTTTGGAGTGAAGGGGAAGCCATTAAATAATTATG  
CCAGCACGGCCGGGTGCGGTGGCTCACGCCTGTAATCCAGCACTTTGGGAGGCCGAGGC  
[T,G]  
GGTGGATCACAGGTCAGGAGATCGAGACCATCCTGGCTAACACGGTGAAACCCCGTCTC

FIGURE 3Q

TACTAAAAATACAAAAATCAGCTGGGCGTGGTGGCGGGCACCTGGAGTCCCAGCTACTC  
 AGGAGGCTGAGGCAGGAGAATGGCGTGAACCCGGGAGGCAGAGCTTGAGTGAGCCAAGG  
 TCACGCCACTGCACTCCAGCCTGGGCGACAGAGTGAGACTCCGTCTCAAAAAATAATAAT  
 TATTATGCCAGCATGGTGGCTCATGCCTATAATCCAGCACTTTGGGAGGCCAAGGCAGG

24558 GAGAGTGAAAGGGGAAGCCATTAATAATTATGCCAGCACGGCCGGGTGCGGTGGCTCAGG  
 CCTGTAATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCACAGGGTCAGGAGATCGAG  
 ACCATCCTGGCTAACACGGTGAAACCCCGTCTCTACTAAAAATACAAAAATCAGCTGGG  
 CGTGGTGGCGGGCACCTGGAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATGGCGTG  
 AACCCGGGAGGCAGAGCTTGAGTGAGCCAAGTACGCCACTGCACTCCAGCCTGGGCG  
 [T, C, A]  
 CAGAGTGAGACTCCGTCTCAAAAAATAATAATTATTATGCCAGCATGGTGGCTCATGCCT  
 ATAATCCCAGCACTTTGGGAGGCCAAGGCAGGATTGCTTGAGGCCAGGAGTTCAAGACCA  
 GCCTGGGCAACATAGCAAGACCCCATCTCTAAAAAAAAAAAAAAAAATTAGCCGGGCGTGGT  
 GTGGGTGCCTGTAGTCCCAGCAACTCAGGAGGCTGAGGTGGGAGGATTGCTTGAGTCTGG  
 GAGGTGGAGGTTGAGTGAGCTGAGATTGCACCACTGTACTCCAGCCTGGGTGACAGAGC

24872 CAAGACCCCATCTCTAAAAAAAAAAAAAAAAATTAGCCGGGCGTGGTGGTGGGTGCCTGTAGT  
 CCCAGCAACTCAGGAGGCTGAGGTGGGAGGATTGCTTGAGTCTGGGAGGTGGAGGTTGCA  
 GTGAGCTGAGATTGCACCACTGTACTCCAGCCTGGGTGACAGAGCCAGACCCTGTCTC  
 [-, A]  
 AAAAAAAAAAGAAAAAAAAAGTAATAATAATTATGCCAGGACAGCAGGTGGACGGACACC  
 TGGTCCTTCTGACTCAGAGCCTGTGGTCCAGCACCCCTAGTGGTGAACAAGCCAGACA  
 CAGGATAAGGATACATTTAGTGTCTAGTTTGTACCTGGCAAACAGAGTGACAAGATTGGG  
 CTTAATACTTTCCAGCTATAAAATTCTAGAATTCTGTGACCCAAGTTAATTTGGGGTAG  
 AGCTTTTTAAAAAAAAATAGAGATGGAGTCTTGCCATGTTGCCCAGGCTGGACTTAAAC

25756 AGCTGGGATTACAGGTGCATGCCACCACCACACCGGCTAATTTTTGTATTTTTATTAGAG  
 ACGGGATTTACCATTTTGGCCAGGCTGGTCTCAAACTCCTGACCTCAGGTAATCTGCCC  
 ACCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTGTGCCCGGCCATAGAGT  
 TTTTATACTTTGGGATAATTGTAGAACTCAGTAGTAGAGTTAAGTGGAGTTGGTCCTT  
 TTTAAAGATATCAAAACCCATTTACTGGTTATTTTAAAAAGAGACATTTTGGGAGGAAAA  
 [C, T]  
 TAGATATAGAAATCTGTTGAATATGTGACAGAATCCCAAGACTGATAGATGGACTCTGCC  
 CTGTGAACAAGGCAAAGAAAAATGCAAAATGAAAGCCTCTCTACCCAGATCTGCTGGGGG  
 ATGACTGAGGTCAACACAGAAGGCCCTCAGGCCGGGCACGGTGGCTCACGCCTGCAATCC  
 CAACACTTTAGGAGGCTGAGGTGGATGGATCGCTTGAGCCCAGGAGTTTGAGACCAGCCT  
 GGGCAACATGGTGAACCCCTGTTTTATAGAGATAAAAAATACAAAAATTAGCTGGGCG

25968 GTAGTAGAGTTAAGTGGAGTTGGTCCTTTTTAAAGATATCAAAACCCATTTACTGGTTAT  
 TTTAAAAAGAGACATTTTGGGAGGAAACTAGATATAGAAATCTGTTGAATATGTGACAG  
 AATCCCAAGACTGATAGATGGACTCTGCCCTGTGAACAAGGCAAAGAAAAATGCAAAATG  
 AAAGCCTCTCTACCCAGATCTGCTGGGGGATGACTGAGGTCAACACAGAAGGCCCTCAGG  
 CCGGGCACGGTGGCTCACGCCTGCAATCCCAACACTTTAGGAGGCTGAGGTGGATGGATC  
 [G, A]  
 CTTGAGCCCAGGAGTTTGAGACCAGCCTGGGCAACATGGTGAAACCCCTGTTTTTATAGAG  
 ATAAAAAATACAAAAATTAGCTGGGCGTGGTGGCATGTGCCTGTAGTCTCAGCTACTCA  
 GGAGGCTGAGGTGGGAGGATCGCTTGAGCCTGGAAGGCAGAGGTTGCAATGAGCTGAGAT  
 TGCACCACTGCACTGCAGCCTGCACGACAGAGCGAGACGCTGTCTCAAAACAACAACAAA  
 ACCACACACACAGAGAGAAGGCCCTTGATTAGGCTGATAGTTGGAGGATGTAGGGAAGTC

26537 TTAGGCTGATAGTTGGAGGATGTAGGGAAGTCAGCTGGGTGAGCTGTGAGCAGCTCCAG  
 AGGCCGTGCTGGGAGGTTTAGACTTCATCTCTGGTCAATGGGGGGCCACGGAGGCGTTGC  
 GGGCTGAGACTGGGGCTGAGAGACCGGCAAGGAGCAACTGCCGTGATGTAGGGAGGCCA  
 GAGGGAGGCCAAGCTTGGGCGAGTGGGTGAAGGGGGCTTTGAGAGATGTGGGATTAGAT  
 TCCTGTGTGTGAGGGAGAGTGTCTCCCTGAGTGATATTCTGACCCTGAGGTCCCTCT  
 [G, C]  
 TCCCTGGTGTCCCCTGAACAGGAAAGAAGGGTTTGGTGAAGGGGGCAAGAGCCCAGAGCT  
 CCCCGGCGTCCAGGAAGACGAGGCTGCATCCTGAGCCCCTGCATGCACCCAGGGCCACCC

FIGURE 3R

GGCAGCACTCATCCCGCGCCTCCAGAGGCCACCCCCTCATGCAACAGCCGCCCCCGC  
AGGCAGGGGGCTGGGGACTGCAGCCCCACTCCCGCCCCTCCCCATCGTGCTGCATGACC  
TCCACGCACGCACGTCCAGGGACAGACTGGAATGTATGTCATTTGGGGTCTTGGGGGCAG

28204

CCTGGAAACATGGATGGACAAGGGCTTTTGGCCACAGGTGTGGGTGTCTGTGGAGGAG  
GGCTTGTTTGGAGAAGGGAGGCTGGCTGGGGGAGAAACCCGGATCCCGCTGCATCTCCGC  
GCCTGTGGGTGCATGTCGCGTGCTCATCTGTTGCACACAGCTCACTCGTATGTCCTGCAC  
TGGTACATGCATCTGTAATACAGTTTCTACGTCTATTTAAGGCTAGGAGCCGAATGTGCC  
CCATTGTCAGTGGGTCCACGTTTCTCCCGGCTCCTCTGGGCTAAGGCAGTGTGGCCCGA  
[C,T,A,G]  
GCTTAAAAAGTTACTCGGTACTGTTTTTAAGAACAACCTTTATAGAGTTAGTGAAGGCAA  
GTTAAGAGCCAATCACTGATCCCCAAGTGTTTCTTGAGCATCTGGTCTGGGGGGACCACT  
TTGATCGGACCCACCCTTGAAAGCTCAGGGGTAGGCCCAGGTGGGATGCTCACCCTGTC  
ACTGAGGGTTTTGGTTGGCATCGTTGTTTTGAATGTAGCACAAGCGATGAGCAAACCTCT  
ATAAGAGTGTTTTAAAAATTAACTTCCAGGAAGTGAGTTAAAAACAATAAAAGCCCTTT

FIGURE 3S